CE LIBRARY
AUG 6 1925

Vol. XVII. Parts I and II June, 1925

BIOMETRIKA

A JOURNAL FOR THE STATISTICAL STUDY OF BIOLOGICAL PROBLEMS

FOUNDED BY

W. F. R. WELDON, FRANCIS GALTON AND KARL PEARSON

EDITED BY

KARL PEARSON

ASSISTED BY

EGON S. PEARSON

ISSUED BY THE BIOMETRIC LABORATORY
UNIVERSITY COLLEGE, LONDON
AND PRINTED AT THE
UNIVERSITY PRESS, CAMBRIDGE

Price Thirty Two Shillings net including postage

PRINTED IN GREAT BRITAIN

[*Issued* 10 *July* 1925]

UNIVERSITY OF LONDON, UNIVERSITY COLLEGE

DEPARTMENT OF APPLIED STATISTICS

The Biometric Laboratory

(assisted by a grant from the Worshipful Company of Drapers)

Until the phenomena of any branch of knowledge have been submitted to measurement and number it cannot assume the status and dignity of a science. FRANCIS GALTON.

Under the direction of Professor KARL PEARSON, F.R.S. Assistants: E. S. PEARSON, M.A., M. MOUL; Hon. Research Assistants: G. M. MORANT, M.Sc., Dr JULIA BELL and KATHERINE WATSON, D.Sc. Crewdson-Benington Student: B. HOOKE.

This laboratory provides a complete training in modern statistical methods and is especially arranged so as to assist research workers engaged on biometric problems.

The Francis Galton Eugenics Laboratory

National Eugenics is the study of agencies under social control, that may improve or impair the racial qualities of future generations, either physically or mentally.

The Laboratory was founded by Sir FRANCIS GALTON and is under the supervision of Professor KARL PEARSON, F.R.S. Assistant Professor: ETHEL M. ELDERTON. Reader in Medical Statistics: M. GREENWOOD, M.R.C.P., M.R.C.S. Medical Officer: PERCY STOCKS, M.A., M.D. Assistants: M. NOEL KARN and J. O. IRWIN, M.Sc. Secretary: M. B. CHILD.

It was the intention of the Founder, that the Laboratory should serve (i) as a storehouse of statistical material bearing on the mental and physical conditions in man, and the relation of these conditions to inheritance and environment; (ii) as a centre for the publication or other form of distribution of information concerning National Eugenics; (iii) as a school for training and assisting research-workers in special problems in Eugenics.

Now Ready, Cambridge University Press

New Series, TRACTS FOR COMPUTERS

This series will endeavour to supply a gap in statistical literature, namely "first aid" to the professional computer.

I. **Tables of the Digamma and Trigamma Functions.** By ELEANOR PAIRMAN, M.A. *Price 3s. net.*

Tables for summing $S = \sum_{i=1}^{\infty} \frac{1}{(p_1 i + q_1)(p_2 i + q_2) \ldots (p_n i + q_n)}$ where the p's and q's are numerical factors.

II. **On the Construction of Tables and on Interpolation. Part I. Univariate Tables.** By KARL PEARSON, F.R.S. *Price 3s. 9d. net.*

III. **On the Construction of Tables and on Interpolation. Part II. Bivariate Tables.** By KARL PEARSON, F.R.S. *Price 3s. 9d. net.*

IV. **Tables of the Logarithms of the Complete Γ-function to Twelve Figures.** Reprint with Portrait of A. M. LEGENDRE. *Price 3s. 9d. net.*

V. **Table of Coefficients of Everett's Central-Difference Interpolation Formula.** By A. J. THOMPSON, B.Sc. *Price 3s. 9d. net.*

VI. **Smoothing.** By E. C. RHODES, M.Sc. *Price 3s. 9d. net.*

VII. **The numerical Evaluations of the Incomplete B-Function or of the Integral** $\int_0^x x^{p-1}(1-x)^{q-1}\,dx$ **for Ranges of x between 0 and 1.** By H. E. SOPER, M.A. *Price 3s. 9d. net.*

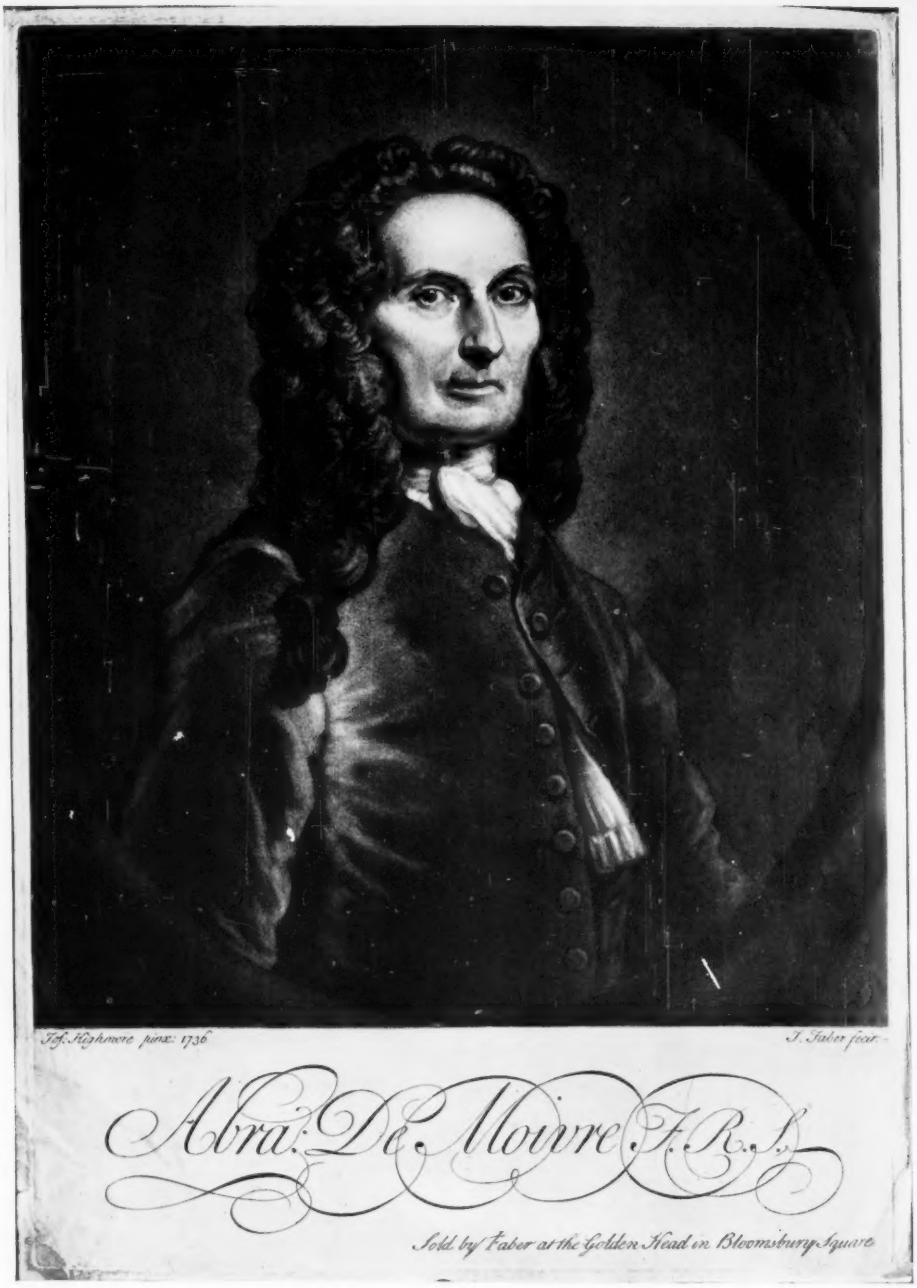
VIII. **Table of the Logarithms of the Complete Γ-Function (to ten decimal places) for Argument 2 to 1200 beyond Legendre's Range (Argument 1 to 2).** By EGON S. PEARSON, M.A. *Price 3s. 9d. net.*

IX. **Log Γ(x) from $x = 1$ to 50·9 by intervals of 0·01.** By JOHN BROWNLEE, M.D., D.Sc. *Price 3s. 9d. net.*

X. **On Quadrature and Cubature or on Methods of Determining Approximately Single and Double Integrals.** By J. O. IRWIN, M.Sc. *Double Number, Price 7s. 6d. net.*

XI. **Logarithmetica Britannica.** A Standard Table of Logarithms to Twenty Decimal Places. Part IX. Logarithms of Numbers 90,000 to 100,000. By A. J. THOMPSON. *Price 15s. net.*

XII. **Tables of the Probable Error of the Coefficient of Correlation.** By KARL HOLZINGER, PH.D. *Price 3s. 9d. net.*



Abraham De Moivre: Discoverer of the "Normal Curve."

Erratum.

Subscribers to *Biometrika* are requested to make the following correction in the paper by Mr Frank Sandon in *Biometrika*, Vol. XVI. upon the "Simplification of the Calculation of Abruptness Coefficients," p. 194, l. 11. In the value for μ_2' a negative sign has been dropped before the term $\frac{2x_p'k_1}{N}$ in the square brackets. A computer recently wasted much time by not observing that the sign of this term was in error.

VOLUME XVII JUNE, 1925 NOS. 1 AND 2

BIOMETRIKA

A STUDY OF EGYPTIAN CRANIOLOGY FROM PREHISTORIC TO ROMAN TIMES.

BY G. M. MORANT, M.Sc., Crewdson-Benington Student in Craniometry.

Together with a series of measurements on Crania of the First Dynasty from Royal Tombs at Abydos. By G. H. MOTLEY, B.Sc.

CONTENTS.

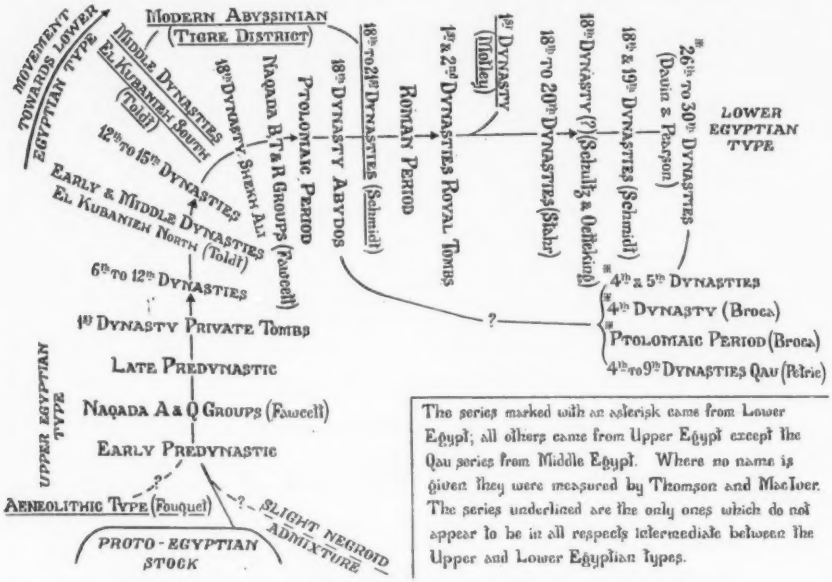
PAGE

(1) Introductory 1
(2) Methods of Measurement and Remarks on the Cranial Anomalies of First Dynasty Royal Tombs skulls from Abydos . . . 10
(3) Prehistoric Series of Egyptian Skulls 12
(4) Skulls of the First to the Sixth Dynasties 22
(5) Series from Lower Egypt later than the Fifth Dynasty . . . 29
(6) Upper Egyptian Series later than the Fifth Dynasty measured by Thomson and MacIver 32
(7) Upper Egyptian Series later than the Fifth Dynasty measured by Chantre and Broca 34
(8) Upper Egyptian Series later than the Fifth Dynasty measured by German Craniometricians 35
(9) Type Contours of First Dynasty Skulls from Royal Tombs at Abydos 44

Appendix:
Table of Absolute and Indicial Measurements . . . *Facing* 52

1. *Introductory.* The craniology of the ancient Egyptians has been more intensively studied than that of any other race of man and perhaps no other people has been accredited at different times with such varied origins. Since Morton (1), who supposed them the posterity of Ham, furnished descriptions and measurements of dynastic crania, the craniometric material has been slowly accumulating, although only a few of the series dealt with in more recent years can satisfy all the requirements of biometry. The present paper provides a notable addition to our knowledge in dealing with hitherto undiscussed measurements of two series of crania which are preserved in the Biometric Laboratory. One is of 50 ♂ and ♀ skulls from Royal Tombs of the First Dynasty at Abydos, measured by Mr G. H. Motley, and the other of 900 ♂ skulls from Gizeh belonging to the 26th to 30th Dynasties, measured by Miss A. Davin and other workers in the Laboratory. Both series were excavated by Sir Flinders Petrie. The latter is the only one available which is long enough to give really adequate representations of a statistical population, but there are several others sufficiently long to give quite reliable mean measurements. An attempt has been made to collect from published sources the mean measurements of all series of Egyptian ♂ crania from prehistoric

to Roman times which contain more than 20 specimens*. In estimating the inter-relations of such abundant material the need for the greatest possible refinement in method is at once felt. The differences between the various types are in many cases so small that it is impossible to determine their affinities from a comparison of isolated characters; the true sequences are obscured by errors of random sampling. In such cases the method of the Coefficient of Racial Likeness, which was originally suggested by Professor Karl Pearson and of which extensive use has now been made, becomes an indispensable aid. That criterion takes cognisance at one time of all the principal measurements of the skull and it gives a measure of the relationship between two series compared as a single numerical quantity.


Modern Abyssinian (Tigre District)
Movement towards Lower Egyptian Type
Middle Dynasties El Kubanieh South (Toldt)
12th to 15th Dynasties
Early & Middle Dynasties El Kubanieh North (Toldt)
6th to 12th Dynasties
1st Dynasty Private Tombs
Late Predynastic
Naqada A & Q Groups (Fawcett)
Early Predynastic
Upper Egyptian Type
Aeneolithic Type (Fouquet)
Proto-Egyptian Stock
? Slight Negroid Admixture
18th Dynasty Sheikh Ali
Naqada B, T & R Groups (Fawcett)
Ptolomaic Period
18th Dynasty Abydos
18th to 24th Dynasties (Schmidt)
Roman Period
1st & 2nd Dynasties Royal Tombs
1st Dynasty (Molley)
18th to 20th Dynasties (Shahr)
18th Dynasty (?) (Schultz & Oehlking)
18th & 19th Dynasties (Schmidt)
26th to 30th Dynasties (Davis & Pearson)
Lower Egyptian Type
4th & 5th Dynasties
4th Dynasty (Broca)
Ptolomaic Period (Broca)
4th to 9th Dynasties Qau (Petrie)
?

The series marked with an asterisk came from Lower Egypt; all others came from Upper Egypt except the Qau series from Middle Egypt. Where no name is given they were measured by Thomson and MacIver. The series underlined are the only ones which do not appear to be in all respects intermediate between the Upper and Lower Egyptian types.

Fig. 1. Diagram of the Inter-relations of Ancient Egyptian Series of Crania.

In calculating this coefficient attention is focussed on the characters which differ most profoundly. Without some such criterion as the Coefficient of Racial Likeness, which can give a mathematical precision to our appreciation of the degree of resemblance between allied cranial types, it would have been almost impossible to determine the true relationship of the numerous ancient Egyptian series and to reduce them to any uniform plan. The comparison of the Coefficients leads to the arrangement of the types shown in Fig. 1.

* It is surprising that so many German craniologists writing in recent years, and with quite a lively sense of the need for full measurements of as many skulls of a homogeneous population as possible, should be content to publish individual measurements without means. I have had to calculate the means of nearly all the material dealt with in this paper with the exception of the series in the Biometric Laboratory.

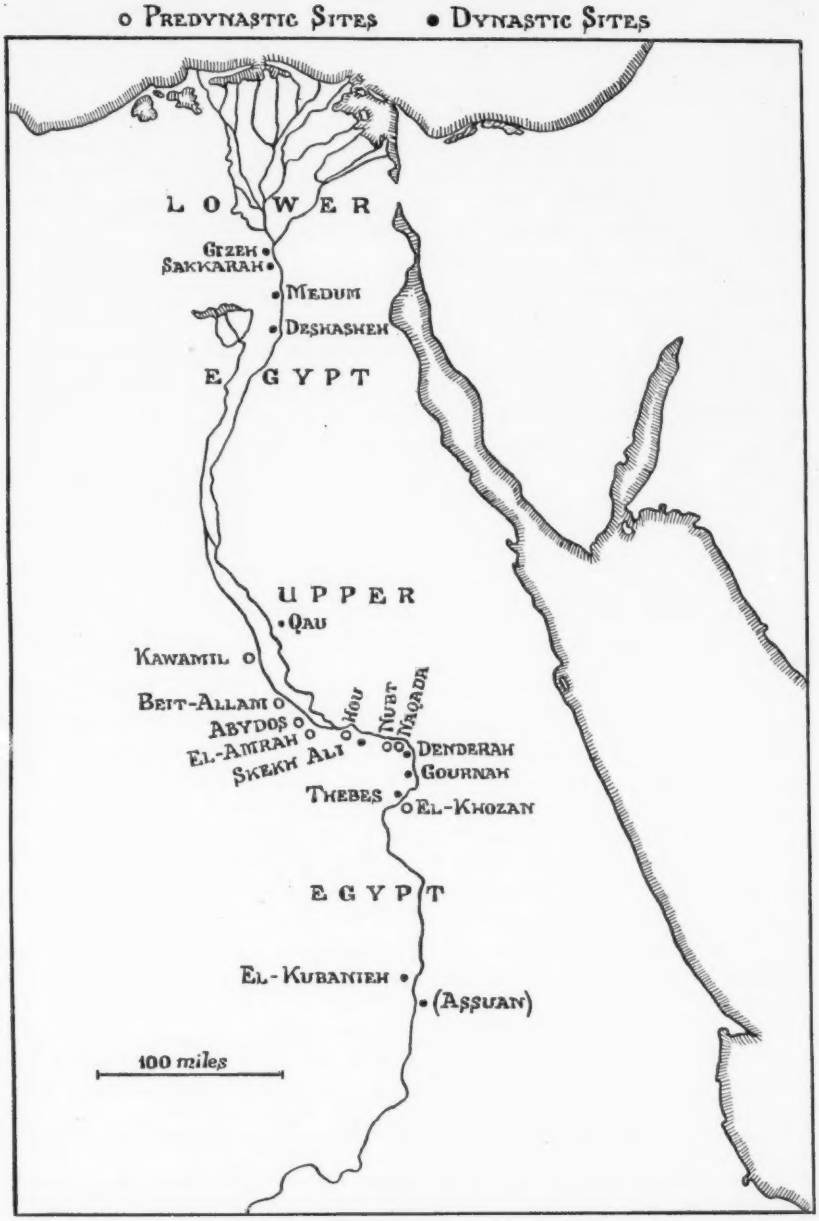


Fig. 2. Map showing sites from which the cranial series dealt with in this paper were obtained.

In his Presidential Address to the Anthropological Section of the British Association in 1923 Professor P. E. Newberry said: "It is remarkable that so little is known about the early history of the Delta. Few excavations have been carried out there, and nothing of Pre-dynastic, or Early Dynastic, times has, so far, been brought to light from the country north of Cairo" ((2), p. 183)*. There are no measurements of Pre-dynastic skulls and very few of Early Dynastic specimens from Lower Egypt, but the long series of 26th to 30th Dynasties skulls from Gizeh makes our knowledge of the physical type prevailing there in later times quite exact. For the district around Thebes and Abydos in Upper Egypt the craniometric material is far more ample than for any other region of the same size in the world.

A detailed comparison of all the ancient Egyptian material leads to the following main conclusions:

(i) In Pre-dynastic times there was a primitive, dolichocephalic race living in Upper Egypt which may have been directly descended from the people whose palaeoliths are found on the high desert floor on either side of the Nile Valley. Though quite distinct from the main population, it bears certain relations to that more advanced type which suggests that the two had a common origin. This so-called Aeneolithic† race may have been assimilated into the main stock of Upper Egypt if its numbers were small, but there is not the slightest suggestion of modification of the skull form of the latter resulting from such an admixture. At any rate the Aeneolithic type had disappeared from the Thebaid before the opening of the historic era.

(ii) In Early Pre-dynastic times there were two distinct races of man living in Egypt; one in the Thebaid and the other, it is supposed, in the Fayum. These may be called the Upper and the Lower Egyptian races. They were as closely related to one another as two adjacent peoples are generally found to be, and there can be no doubt that they diverged from the same branch of the human tree at no very early date. The earliest direct evidence we have of the Lower Egyptian type is in the persons of the royal retainers of the First and Second Dynasties who were buried in Abydos, but it is probable that they were intrusive in Upper Egypt as the type is quite distinct from that of the contemporary native population. The earliest series of Lower Egyptian type found in the Fayum belong to the Fourth Dynasty, but the presumptive evidence in favour of the view that the population there was of almost precisely the same type in much earlier times is very strong. The Upper Egyptian type is known directly from crania found in the Thebaid and dating from Early Pre-dynastic times.

(iii) The Lower Egyptian type seems to have remained unchanged from Early Dynastic to Ptolemaic times except that a relatively small part of the population was modified very slightly, possibly by admixture with some unknown foreign race.

* All the sites from which skull series dealt with in this paper were obtained are shown on the accompanying map (Fig. 2).

† In using the term *Aeneolithic*, as in Italian usage, I am adopting a suggestion made by Sir Flinders Petrie.

(iv) The Upper Egyptian type was slowly transformed from the very earliest times in which we have acquaintance with it, and by Late Dynastic times the population of Upper Egypt was of almost pure Lower Egyptian type.

(v) The pure Upper and Lower Egyptian types bear the following relations to one another:

(*a*) All characters measuring lengths in the sagittal plane, except the basio-alveolar length GL (viz. L, LB, S, S_1, S_2, S_3, S_2', *fml* and G_1'), facial breadth characters (GB, O_1', G_2 and NB), $A\angle$ and $B\angle$ of the fundamental triangle, the profile angle, occipital index and palatal index are identically the same for the two types.

(*b*) The facial heights $G'H$ and NH', foraminal breadth (*fmb*) and upper facial index ($100\,G'H/GB$) of the Lower Egyptian type are slightly greater than those of the Upper, but the height index $100\,H'/L$ is just appreciably smaller.

(*c*) Considerable differences are found between the characters Q', U, J, B' and cranial capacity, all being in this case primarily dependent on differences between the calvarial breadths, and the orbital height and index are greater for the Lower Egyptian type, while, in keeping with the latter difference, the nasal index is less. With its much greater calvarial breadth the Lower Egyptian skull has the smaller basio-bregmatic height and its basio-alveolar length and nasal angle are also smaller than those of the southern type.

(*d*) The greatest differences are between the values of the calvarial breadth, the cephalic index, which is entirely dependent on B as L is constant, and the index $100\,B/H'$, the latter being a peculiarly distinctive character as B and H' bear opposite relations to one another.

We conclude then that the calvaria of the Lower Egyptian skull has a greater breadth, a rather smaller height and the same length as that of the Upper Egyptian type. The facial bones of the two, other than the zygomatic, are equal in breadth but the broader skull has the rather longer face.

These two types represent the extremes of the pure native Egyptian population from Early Pre-dynastic to Ptolemaic times and the relations between them are of the same nature as those which two allied, adjacent and *contemporaneous* races would bear to one another in general. The vast majority of the Dynastic series of crania from Upper Egypt are intermediate between the Upper and Lower Egyptian types. Characters—such as those related to the calvarial length—which are the same for the extremes are constant for all and where the greatest differences are shown the transition from one to the other can be clearly traced. The series represent various stages in the transformation of one type into the other. The types shown in the following table were selected because each represents an adequate number of crania* and shows no signs of intermixture with any race foreign to Egypt. The order in which they are arranged is that deduced from a comparison

* No mean measurement shown in this table is based on fewer than 36 skulls and all those not marked with an asterisk are based on more than 50.

of the Coefficients of Racial Likeness between them. Every series came from Upper Egypt except that of the 26th to 30th Dynasties which represents the pure Lower Egyptian type.

Series	Measured by	*B*	100 *B/L*	100 *B/H'*	*H'*	*J*	100 *NB/NH'*	*N*∠	*GL*	*L*	*LB*	*A*∠
Early Pre-dynastic	Thomson & MacIver	131·4	71·7	98·1*	134·0*	124·3*	50·3*	67°·0*	98·8*	183·5	102·0*	72°·6*
Naqada *A* & *Q* Groups	Fawcett ...	132·7	71·8	99·2	133·8	125·9*	50·0	65°·5*	96·7*	184·7	101·4	73°·3*
Late Pre-dynastic	Thomson & MacIver	133·4	72·1	100·5	132·7	127·0	51·0	66°·8	98·3	185·1	101·6	72°·4
6th to 12th Dynasties	,,	134·3	73·6	99·6	134·8	124·5	50·4	65°·7	96·5	182·5	100·8	72°·8
12th to 15th Dynasties	,,	133·8	73·1	100·6	133·0	126·4	48·8	65°·7	96·7	183·0	100·6	72°·0
Early and Middle Dynasties (El Kubanieh South)	Toldt ...	135·9	74·4	102·3	132·9	125·7	50·8	67°·5	98·1	182·7	100·1	71°·0
Ptolemaic Period	Thomson & MacIver	135·3	73·7	102·9	131·5	128·6	49·3	64°·3	95·0	183·5	100·7	73°·3
18th to 21st Dynasties	E. Schmidt	136·6	75·1	103·6	131·9	128·3	50·1	66°·1	96·4	181·9	100·6	73°·0
Roman Period	Thomson & MacIver	136·4	75·2	104·4	130·7	127·5*	49·4*	64°·3	94·2	181·4	99·9	73°·2
18th to 20th Dynasties	Stahr ...	137·3	74·8	102·7	133·7	128·6	50·4	64°·5	96·1	183·5	101·9	74°·0
26th to 30th Dynasties	Davin and Pearson	138·9	75·1	104·9	132·4	128·7	47·3	63°·9	95·0	185·3	101·6	74°·1

* No mean measurement shown in this table is based on fewer than 36 skulls and all those not marked with an asterisk are based on more than 50.

The differences between adjacent types in the above table are so small that it is surprising to find that the sequences of characters are as regular as they appear to be. If the numbers on which the mean measurements were based had been smaller the orders would very probably have been hopelessly obscured for the characters showing the smaller differences between the extremes. All the characters show significant differences except *L*, *LB* and *A*∠ on the extreme right of the table and the constancy of those three is very striking.

(vi) The question how exactly the population of Upper Egypt was transformed from its pristine type to that of the Lower Egyptian race is one of considerable importance from both an evolutionary and an historical point of view. To answer it with any degree of finality we should require the evidence not of 5000 but of 50,000 skulls and our present conclusions can be no more than tentative. The available series are not scattered randomly in time: there are very few for Pre-dynastic and Early Dynastic times, the later dynasties are well represented but the evidence relating to Ptolemaic and Roman times is meagre. The 18th to 20th Dynasties are far more adequately represented than any other period of the same length*. If the Pre-dynastic period was as long as the Dynastic, as is generally

* That is by the number of cranial series, but if the number of skulls is considered the 26th to 30th Dynasties are far better represented than any other period as the long E series belongs to that time.

supposed, then the change in the Upper Egyptian type prior to the First Dynasty was exceedingly slow, but quite uniform as far as we can tell. Such a conclusion is quite in keeping with the archaeological evidence. In the early dynasties the movement became more rapid and by the 15th the skull form of the natives of Upper Egypt was approximately mid-way between those of the Upper and Lower Egyptian types. Either during or immediately preceding the 18th Dynasty there seems to have been a considerable acceleration of the movement. Four series belonging to the 18th to 20th Dynasties—two from Thebes and two from Abydos—are of almost pure Lower Egyptian type, but like the Royal Tombs series of the First and Second Dynasties, to which they bear the closest possible resemblance, they still incline slightly towards the Upper Egyptian type. It is possible that those five series indicate sudden intrusions into Upper Egypt which were independent of the slower movement*. There are two other 18th and 19th Dynasties series which represent the slower and more normal transformation which pursued its course steadily down to Ptolemaic and Roman times. The rate of change after the First Dynasty appears to have been on the whole remarkably uniform and there is no suggestion of a retrograde movement—that is to say a reversion to the Upper Egyptian type—at any time.

(vii) As examples of types which were entirely unaffected for several thousand years by any influences foreign to their country, the ancient Egyptians may well be unparalleled in the history of the world. All the series which do not appear to be of absolutely pure Egyptian type are underlined in Fig. 1 and, with the exception of Fouquet's Aeneolithic type, the divergences from the main stock are exceedingly small†. The most aberrant series is Motley's First Dynasty from the Royal Tombs of Abydos. It is most similar to the Lower Egyptian type, though inclining slightly towards the Upper Egyptian—the royal retainers being supposed intrusive in the Thebaid—but it also shows signs of admixture with an alien race which was certainly not Negroid. The calvariae of these First Dynasty skulls are not in any way peculiar, but the facial heights, and to a less marked extent the facial breadths, are characteristically large. No other series dealt with in this paper has shown any tendency to diverge in the same direction. Three series from Lower Egypt are of pure Lower Egyptian type for all characters except the basio-bregmatic height which has values which are quite significantly greater than that of the 26th to 30th Dynasties type‡. An 18th Dynasty series from Abydos, which in all other ways is intermediate between the Upper and Lower Egyptian types, is also

* It is quite probable that the 18th Dynasty series which closely resemble the pure Lower Egyptian type were representatives of the middle and upper classes of the population, while the lower social orders were closer to the pure Upper Egyptian type. Thomson and MacIver have provided measurements of two 18th Dynasty series. One from Shekh Ali was found in a pit and it undoubtedly represented an inferior class; the other, from Abydos, was obtained in a middle class cemetery. The Shekh Ali series is distinctly closer to the earlier type of Upper Egypt than the other. (See (13), p. 29.)

† At different times during the Dynastic era there were undoubtedly quite large numbers of foreign slaves in the country. The only skull measurements of such people available seem to be those of pygmy crania of the Third Dynasty (8).

‡ Sir Flinders Petrie's field measurements of Fourth to Ninth Dynasties skulls from Qau indicate a precisely similar divergence (see p. 29, ftn. † below).

characterised by a large basio-bregmatic height. Judging from means reduced by Lee from Schmidt's individual measurements (see (6), p. 426), the modern Copts have diverged greatly from the type of the dynastic Egyptians and in no way more significantly than in their reduced length ($L = 177{\cdot}0$) and greater cephalic index ($100B/L = 77{\cdot}3$). But the modern natives of Northern Abyssinia measured by Sergio Sergi (26) are surprisingly similar to the type which predominated in Upper Egypt during the middle dynasties. The relationship between one of the latter and the Abyssinians is decidedly closer than that between two series—the Whitechapel and Moorfields—of seventeenth century Londoners. This is a striking example of the persistence of a type with only slight modification for a period of at least 3000 years. There are two middle dynastic series which diverge slightly from the pure native population contemporaneous with them in the ways which were more accentuated in the later Abyssinians.

(viii) It is very generally supposed that the population of ancient Egypt was sensibly affected at various times by the infusion of Negro blood. But in the series of which we have the mean measurements it is not possible to detect the slightest effect of any such admixture that can have taken place after early Pre-dynastic times. Apart from isolated negroid skulls which are said to have been excluded from several of the series, the populations appear to be quite homogeneous and we have no reason to suppose that the mean type was affected in the slightest by admixture with any race foreign to Egypt. But it is quite possible that the race we have called the pure Upper Egyptian had originally some Negroid elements in its make-up. There are very few measured series of Negro skulls which are sufficiently long to give reliable determinations of type. A comparison of the more reliable material shows that there is a considerable variation in calvarial measurements among Negro races: all the Egyptian series fall within the extremes of length, breadth and cephalic index shown by the Negro types. But certain other measurements—chiefly facial—apparently distinguish all Negro from all Egyptian types. In the table on p. 9 the Upper Egyptian type is represented by the Early Pre-Dynastic and Naqada series and the Lower Egyptian by the 26th to 30th Dynastic series. The Negro races were selected because more or less adequate skull measurements have been furnished for them but for no other reason. All the measurements are of ♂ skulls.

The characters $NH(R)$, NB, $NB/NH(R)$, $P\angle$, $N\angle$ and $B\angle$ clearly distinguish the Egyptian from all Negroid types: they seem to be very constant for the latter. At the same time we find that for each one of those measurements the Upper Egyptian types are nearer to the Negro values than are the Lower Egyptian, and the Early Pre-dynastic is very slightly nearer to the Negroes than is the Naqada type. Much the same relations are found for the basio-alveolar length (GL), although the low value for the Congo series is discordant, and the indices $100H'/L$ of the Upper Egyptian types are also between the Lower Egyptian and Negro values. With the facial height ($G'H$) and index ($100G'H/GB$) the case is rather different. For those two characters the Bantu Negroes are almost identically the same as the ancient Egyptians, but the smaller values of the non-Bantu races differentiate that

		Number of Skulls	*NH, R*	*NB*	100 *NB/NH, R*	*P* ∠	*N* ∠	*B* ∠	*GL*	*G'H*	100 *G'H/GB*	100 *H'/L*	S_3'	S_3	*Oc. I.*
Ancient Egyptians	26th to 30th Dynasties	850	51·7	24·4	47·3	85°·8	63°·9	41°·9	95·0	70·4	74·1	71·5	94·7	115·2	61·5
	Naqada *A* and *Q* Series	70	50·2	25·0	49·8	84°·6	65°·5	41°·2	96·7	69·7	72·7	72·4	96·5	116·9	60·2
	Early Predynastic	40	50·1	25·2	50·3	—	67°·0	40°·4	98·8	69·2	—	73·3	—	—	—
Bantu Negroes	Zulu* ...	20	47·2	27·3	58·1	—	70°·1	39°·8	101·9	69·5	72·1	75·0	—	111·7	—
	Angoni* ...	25	48·5	28·0	58·0	—	70°·5	39°·9	103·5	70·3	74·1	74·3	—	108·9	—
	Kaffir* ...	35	48·6	27·3	56·3	—	70°·7	38°·6	105·1	69·1	72·2	72·1	—	119·3	—
Non-Bantu Negroes	Congo* ...	50	47·2	26·0	55·2	82°·8	69°·3	37°·9	96·5	63·4	66·9	75·4	95·0	107·7	68·8
	Gaboon, 1864 Series*	50	48·2	26·6	55·3	—	70°·5	38°·6	100·5	66·4	69·7	75·5	95·9	109·1	68·3
	Northern Negro* ...	30	49·5	27·3	55·4	82°·0	72°·4	38°·0	105·3	68·0	70·4	74·2	—	—	—
	Cameroons† .	80	49·1	27·5	56·0	80°·5	72°·0	39°·1	102·6	67·9	68·1	75·0	97·6	112·4	65·6
	East Coast (Dschagga)‡	12	44·3	27·8	62·2	—	73°·0	38°·6	105·0	68·2	70·5	73·2	—	—	—

* The series marked with an * were taken from (7), pp. 298, 299. The Zulu, Angoni and Kaffir means were reduced from individual measurements provided by Shrubsall. The Congo and Gaboon series were measured by Crewdson-Benington and the Northern Negroes were taken from various parts of the German Anthropological Catalogue (19). Of the last series it was said it can "only be of service when we are comparing negroes in general with other races": see (7), p. 295.

† The means of these Cameroons' skulls were reduced from individual measurements provided by Drontschilow (25). In calculating them I omitted 7 pathological and 3 young skulls.

‡ The individual measurements of these Dschagga skulls are provided by Widenmann (24).

group from the others although they are more similar to the Upper than to the Lower Egyptian values. The flattened occipital bone of the non-Bantu Negro skull is a most characteristic feature but we cannot say that in that respect one Egyptian race is more negroid than the other. From a comparison of facial measurements alone we can at least say that there is a distinct suggestion of negroid affinities in what has been called the Upper Egyptian race. The calvarial measurements do not support that view, and we cannot gauge the significance of the observed relationships while any knowledge relating to the skull form of a type which has resulted from the crossing of two races as markedly different as a pure Negro and a pure Egyptian is quite lacking.

In dealing with the long-bones of the Naqada series Warren (9) found that there was a suggestion of negroid affinities*.

(ix) A study of all the series of early Egyptian crania that have as yet been measured shows that they are inter-related in a simple way. It is most probable that the observed differences between the types are to be explained as being due to some slow moving evolutionary process—whether through selection or directly through environment—or to a slow blood mingling of different races. On the latter hypothesis it would only be necessary to assume the existence of two original races, an Upper and a Lower Egyptian, which were distinct in Early Pre-dynastic times. The gradual approach of the Upper Egyptian population to the Lower Egyptian type—a movement, it should be noted, which was not completed as late as Roman times—would then be due to admixture in the Thebaid, while the original northern type was apparently persisting in the Fayum without modification. On the other hand, an evolutionary hypothesis would give an explanation of the facts which would be equally reasonable. Additional evidence, however, such as that which would result from a study of the variabilities of the series dealt with in this paper or the discovery of an adequately long Pre-dynastic cranial series in Lower Egypt, might tell strongly in favour of the one hypothesis or the other. It may be that the observed changes were the result of both racial admixture and an evolutionary process.

2. *Methods of Measurement and Remarks on the Cranial Anomalies of First Dynasty Royal Tombs skulls from Abydos.* The individual measurements taken by Mr G. H. Motley and "remarks" on the more outstanding cranial anomalies are contained in the Appendix to this paper. The methods of measurement used in dealing with the First Dynasty and the 26th to 30th Dynasties series from Gizeh, which was measured by Miss A. Davin and others, are precisely those which are ordinarily taken in the Biometric Laboratory†. The majority are in accordance

* Warren says: "On the whole, the proportions of the limb-bones to one another may be said to have approached those of the Negro, while the sacral and scapular indices were almost identical with those of Europeans." ((9), p. 191.)

† Full descriptions of the measurements are given in *Biometrika*, Vol. xiv. pp. 196—200. Briefly: F=Ophryo-occipital length, L'=glabellar projective length, L=glabella-occipital length, B=maximum parietal breadth, B'=least forehead breadth, H'=basio-bregmatic height, H=Frankfurt vertical height, OH=Frankfurt auricular height, LB=basion to nasion, Q=transverse arc through "apex" terminating at ear-rods of craniophor, Q'=similar to Q terminating at auricular points, S=sagittal arc

with those of the *Frankfurter Verständigung*. Two additional measurements are given for the First Dynasty and for the re-measured Naqada skulls (see notes to Table I). One is the measurement of nasal height (*NH'*) from the nasion to the base of the anterior nasal spine which is used by Broca, Flower and the majority of French and English craniometricians. *NH'* is slightly less than the Frankfurt measurements *NH, R* and *NH, L*. The other (*EOW*) may be called the external ocular distance: it is taken between the points where the borders of the orbital ridges right and left meet the fronto-malar sutures (see pp. 22 and 23 below). *EOW* is the same as Martin's *innere orbitale Gesichtsbreite**: that and the *äussere orbitale Gesichtsbreite* are occasionally taken by German craniometricians.

Remarks on the individual crania of the First Dynasty series are given in the Appendix. Unless otherwise stated each skull is complete with a mandible and of adult age. The following characteristics and anomalies were looked for.

Signs of age. Condition of sutures, development and loss of teeth and the falling in or thinning of the calvaria.

Condition of teeth. The teeth of this series were remarkably well preserved though all were worn. Many of the skulls that were ageing had not lost a single tooth. Hardly any cases of irregular formations round the palate were found and there was only one carious tooth.

Palate. Of the 48 skulls only one has bony ridges across the palatine grooves leading from the pterygo-palatine canals. This is an unusually low percentage but it is quite possible that ridges originally possessed by other specimens had been broken.

Pre-condyles. No pre-condyles were found at the basion or opisthion but there were several cases of projecting spines of bone in other regions.

Tympanic perforation. All cases were recorded.

Base of the pyriform aperture. This feature is indicated by the letters P.B. The floor is described as "flat"—i.e. parallel to the Frankfurt horizontal plane—"sloping upwards," meaning sloping upwards towards the pyriform border, or

nasion to opisthion, S_1 = arc nasion to bregma, S_2 = arc bregma to lambda, S_3 = arc lambda to opisthion, S_3' = chord lambda to opisthion, U = horizontal circumference, PH = alveolar point to tip of anterior nasal spine, $G'H$ = nasion to alveolar point, GB = breadth between lowest points on zygomatic-maxillary sutures, J = zygomatic breadth, NH, R and L = nasion to lowest edge of pyriform aperture R and L, NB = nasal breadth, DS = dacryal subtense, DC = dacryal chord, DA = dacryal arc, SS = simotic subtense, SC = simotic chord, O_1, R and L = breadths of orbits using Fawcett's curvature method, O_1' = orbital breadth from dacryon, O_2, R and L = heights of orbits, G_1 = palate length from point of *spina nasalis posterior* to an imaginary line tangential to the inner rims of the alveoli of the middle incisors, G_1' = similar to G, from base of spine, G_2 = palate breadth between inner alveolar walls at second molars, GL = basion to alveolar point, *fml* = length of *foramen magnum*, *fmb* = breadth of *foramen magnum*, $P\angle$ = Frankfurt profile angle, *Oc. I.* = occipital index, $N\angle$, $A\angle$ and $B\angle$ found from lengths LB, $G'H$ and GL with aid of Pearson's Trigonometer, $\theta_1 = 180° - P\angle - N\angle$, $\theta_2 = P\angle - A\angle$, EH and EB = height and breadth of palate using Pearson's Uraniscometer. The capacities were found by the method described by Macdonell (*Biometrika*, Vol. III. pp. 203—206).

* Rudolf Martin, *Lehrbuch der Anthropologie*, p. 552, Ed. 1914.

"sloping downwards." The edge is said to be sharp, flat or rounded. Excluding the Negro skull, the following frequencies are found for the combined ♂ and ♀ First Dynasty Egyptians.

		Edge of Pyriform Aperture			Totals
		Sharp	Blunt	Rounded	
Floor of Pyriform Aperture	Sloping upwards ...	12	4	0	16
	Flat...	9	7	2	18
	Sloping downwards.	5	2	2	9
	Totals ...	26	13	4	43

Asymmetry. *SL* or *SR* denotes whether the Sylvian depression is greater on the right or left side and in the same way *JR* and *JL* refer to the relative sizes of the jugular foramina. The following frequencies are found and the balance in favour of the right side is similar to that usually found in racial series.

	Sylvian Depression	Jugular Foramen
Greater on right side	14	20
Equal	28	10
Greater on left side...	3	10

There were few cases of calvarial or facial asymmetry.

Malar bones. No horizontal sutures across the malar bones were found.

Metopism. There were no cases of the persistent frontal suture.

Ossicles in the sutures. These, as ordinarily, were confined almost entirely to the lambdoid suture. There was one ossicle of the λ.

Conformation at the pterion. All cases of epipteric bones and processes of the temporal articulating with the frontal were recorded.

Interparietal bones. There was one case of the simple interparietal and one of the tri-partite interparietal with the *os pentagonale* fused into the supra-occipital.

Other features. Cases of flattening of the obelion are recorded. In a number of the skulls a slight post-coronal depression was observed, but it was not very marked in any case. These skulls seem to be more symmetrical and to exhibit fewer anomalies than most racial series.

3. *Prehistoric Series of Egyptian Skulls.* The earliest series of prehistoric Egyptian crania of any length, and of which measurements have been provided, belong to a period in which metal was used. Palaeolithic implements have been found in abundance on the terraces of the Nile, but it may be doubted whether the country ever passed through a true Neolithic phase. The opinion of Sir Flinders Petrie ((4), p. 28) is that as soon as the sediment of the river had been deposited to a sufficient depth to make agriculture possible, an alien people—supposed Libyan—who were already acquainted with the use of copper, invaded the country and settled along the banks of the Nile. The autochthonous population at that time appears to have been one of palaeolithic hunters and the

invaders did not exterminate either the people themselves or their primitive culture. It is unfortunately not yet possible to determine with any great accuracy the chronological sequence of the various Pre-dynastic series of crania that are available. In the case of the prevailing type an order can be given by using the method of Petrie's sequence dates (see (4), pp. 4—12) which are determined chiefly from a comparative study of the pottery; but it is apparently not possible to establish with certainty the relative ages of the burials of the primitive peoples who maintained their lower stage of culture though living in close proximity to the invaders. All the skulls that can be dealt with came from a comparatively small area of Upper Egypt round Thebes and Abydos. There was another centre of civilisation in Pre-dynastic times in the Fayum*, but no measurements of early skulls from that region are available and there are none of the contemporaneous populations of Middle Egypt.

The affinities of the series of skulls may be estimated by comparing the Coefficients of Racial Likeness† between their mean measurements which are shown in Table I (p. 14). In the majority of cases the numbers of skulls on which the means are based are fairly adequate, but there are few series for which more than half of the 31 characters which should ideally be used have been provided. It is important to remember that a generalised measure of relationship which does not take into account several important regions of the skulls may be misleading and particular caution has to be taken when comparing such measures with those calculated for all the 31 characters.

In the appendix to the first volume of J. de Morgan's researches into the origin of Egyptian civilisation (10) Dr Fouquet gave individual measurements of 11 skulls that had been disinterred at El-Amrah near Abydos‡. From the contracted position of the bodies and the nature of the grave furniture, which comprised abundant pottery but very little metal, it was reasonable to infer that the burials were of Pre-dynastic date though from the description of the artefacts that has been given it is not possible to assign them to any particular period of that era. The mean measurements of the nine specimens are given in the second column of our Table I. The series in columns 3, 4 and 5 of that table are the longest provided by Fouquet§ in the second volume of de Morgan's researches (11) and they were thought to be contemporary with the El-Amrah skulls. It is clear that the four series represent very similar if not identical types. The numbers of skulls in each series are small but the method of the Coefficient of Racial Likeness may profitably be applied. The Coefficients in Table II (p. 16) were calculated for 19 characters of which six were indices. The numbers in brackets are the mean numbers of skulls available for those 19 characters in the case of each series.

* A map showing all the principal Pre-dynastic stations in Egypt is given by J. de Morgan ((10), p. 68).

† See *Biometrika*, Vol. XVI. pp. 11—14 for a definition of the Coefficient of Racial Likeness and an account of the precautions that have to be taken when applying it in practice.

‡ All sites referred to in this paper are shown on the accompanying map (Fig. 2, p. 3).

§ Two series from Naqada North and Guebel-Silsileh ((11), pp. 326 and 336 respectively) contain fewer than five skulls each and their means are not sufficiently like those of the longer series to make it safe to pool them.

TABLE I.

Mean Measurements of Pre-dynastic Egyptian Series. Males.

Pre-dynastic Date	Early ?	Early ?	Early?	Early?	Early?	Middle?	?	Early	Late	Late?‖
Locality... ...	El-Amrah	Kawamil	Beit-Allam	Naqada South	Kawamil, Beit-Allam and Naqada, South	Naqada (*A* and *Q* Series)	Naqada (*B*, *T*, and *R* Series)	Abydos, El-Amrah and Hou	El-Amrah and Hou	El-Khozan
Measured by	Fouquet	Fouquet	Fouquet	Fouquet	Fouquet	Fawcett and Morant	Fawcett and Morant	Thomson and MacIver	Thomson and MacIver	Chantre
F	—	—	—	—	—	184·1 (71)	182·9 (21)	182·6 (53)	183·5 (127)	—
L	189·2 (9)	187·9 (19)	189·3 (13)	188·2 (28)*	188·3 (60)	184·7 (101)	183·9 (24)	183·5 (54)	185·1 (127)	182 (24)
B	137·6 (9)	138·9 (18)	134·2 (13)	137·3 (28)	137·1 (59)	132·7 (88)	136·1 (22)	131·4 (53)	133·4 (125)	133 (24)
H'	137·0 (8)	137·1 (16)	138·5 (7)	136·9 (27)	137·2 (50)	133·8 (70)	132·2 (18)	134·0 (40)	132·7 (107)	130 (24)
B'	93·6 (9)	96·3 (19)	98·2 (13)	98·3 (28)	97·6 (60)	92·6 (101)	91·8 (21)	—	—	92 (24)
S_1	132·1 (9)	129·2 (19)	128·7 (13)	125·9 (28)	127·6 (60)	127·7 (100)	127·4 (22)	—	—	—
S_2	136·2 (9)	132·3 (19)	135·7 (13)	134·1 (28)	133·9 (60)	130·2 (95)	129·7 (25)	—	—	—
U	512·6 (9)	520·8 (18)	516·1 (11)	515·2 (28)	517·1 (57)	510·4 (90)	512·3 (18)	—	—	—
LB	103·4 (9)	106·1 (12)	105·3 (6)	105·2 (27)	105·5 (45)	101·4 (69)	100·1 (18)	102·0 (36)	101·6 (99)	—
G'H	—	—	—	—	—	69·7 (52)	69·6 (18)	69·2 (39)	70·0 (97)	—
GL	—	—	—	—	—	96·7 (46)	94·3 (16)	98·8 (36)	98·3 (97)	—
J	135·6 (5)	130·0 (1)	132·0 (4)	132·3 (18)	132·1 (23)	125·9 (37)	126·2 (15)	124·3 (34)	127·0 (87)	121 (24)
NH'	53·3 (6)	52·3 (12)	51·2 (8)	49·7 (23)	50·7 (43)	50·0 (54)	50·3 (19)	50·1 (37)	50·1 (102)	48 (24)
NB	24·3 (7)	25·7 (12)	24·4 (7)	25·4 (23)	25·3 (42)	25·0 (64)	25·1 (18)	25·2 (37)	25·5 (102)	25 (24)
DC	20·7 (7)	21·1 (12)	20·5 (7)	22·2 (22)	21·6 (41)	—	—	—	—	24 (24)
O_2	34·2 (7)	33·2 (13)	33·9 (8)	32·5 (26)	32·9 (47)	32·2 (57)	32·1 (19)	—	—	32 (24)
O_1'	38·5 (7)	39·2 (13)	39·3 (8)	39·9 (26)	39·6 (47)	39·1 (57)	39·1 (19)	—	—	38 (24)
fml	34·9 (8)	37·1 (14)‡	39·0 (7)	36·0 (27)	36·8 (48)	35·6 (70)	35·5 (19)	—	—	34 (24)
fmb	29·8 (8)	30·1 (14)	29·9 (8)	29·8 (26)	29·9 (48)	29·5 (70)	29·7 (18)	—	—	28 (24)
100 *B*/*L*	72·8 (9)	73·1 (18)	70·6 (13)	72·7 (28)	{72·8 (59)}	{71·8 (88)}	{74·0 (22)}	{71·7 (53)}	{72·1 (125)}	73·4 (23)
100 *H'*/*L*	{72·4 (8)}	72·3 (16)	73·9 (7)	72·8 (27)	{72·9 (50)}	{72·4 (70)}	{71·9 (18)}	{73·3 (40)}	{71·8 (105)}	{71·4 (24)}
100 *B*/*H'*	{100·3 (8)}	{101·3 (16)}	{96·9 (7)}	{100·3 (27)}	{99·9 (50)}	{99·2 (70)}	{103·0 (18)}	{98·1 (40)}	{100·5 (107)}	{102·3 (24)}
100 (*B*−*H'*)/*L*	{0·3 (8)}	{1·0 (16)}	{−2·3 (7)}	{0·2 (27)}	{−0·1 (50)}	{−0·6 (70)}	{2·1 (18)}	{−1·6 (40)}	{0·3 (105)}	{1·6 (24)}
100 *NB*/*NH'*	{45·6 (6)}	49·0 (12)	48·5 (7)	51·7 (23)	{49·9 (42)}	{50·0 (54)}	{49·9 (18)}	{50·3 (37)}	{51·0 (101)}	{52·1 (24)}
100 O_2/O_1'	{88·8 (7)}	85·4 (13)§	85·8 (8)	82·1 (26)	{83·1 (47)}	{82·4 (57)}	{82·1 (19)}	—	—	{84·2 (24)}
100 *fmb*/*fml*	{85·4 (8)}	82·3 (13)	77·6 (5)	82·7 (26)†	{81·3 (48)}	{82·9 (70)}	{83·7 (18)}	—	—	{82·4 (24)}
N ∠	—	—	—	—	—	{65°·5 (46)}	{64°·3 (16)}	{67°·0 (36)}	{66°·8 (97)}	—
A ∠	—	—	—	—	—	{73°·3 (46)}	{73°·9 (16)}	{72°·6 (36)}	{72°·4 (97)}	—
B ∠	—	—	—	—	—	{41°·2 (46)}	{41°·8 (16)}	{40°·0 (36)}	{40°·8 (97)}	—
Capacity	[1473·3 (8)]	[1476·4 (16)]	[1460·5 (7)]	[1465·6 (27)]	[1466·8 (50)]	1388·6 (61)	1382·8 (18)	[1384·0 (40)]	[1396·2 (107)]	[1361·6 (24)]

Notes and footnotes to Table I.

The indices and angles in curled brackets were determined from the mean measurements of the component lengths instead of from the individual measurements. The capacities in square brackets were reconstructed by the formula given by Lee and Pearson (*Phil. Trans.* Series A, Vol. 196, p. 247): $C = \cdot000266 \times L \times B \times H + 524{\cdot}6$ c.cs. H' was used in place of the slightly larger H, so the calculated capacities are likely to be rather smaller than the true ones.

Through the kindness of Professor Karl Pearson, I was able to examine the Naqada crania dealt with by Fawcett(6) and to take several measurements that had not been previously found. When first brought to England the skulls were in a very fragile state and it is no longer possible to take any measurements of several that were included in Fawcett's tables of individual measurements. Excluding those skulls, I found a close agreement between Fawcett's measurements and my own for all characters except the following (see pp. 18—19 below):

	B	B'	LB	$G'H$	GL	NH, R	G_1	O_1R
Fawcett	134·5 (122)	91·2 (130)	99·9 (100)	67·3 (80)	95·0 (68)	49·1 (84)	55·0 (74)	42·5 (82)
Morant	133·4 (110)	92·5 (122)	101·1 (87)	69·7 (70)	96·1 (62)	50·3 (73)	52·0 (66)	41·6 (76)

The disparity between the numbers of skulls on which the means are based cannot account for the whole of the differences between them and we are bound to accept the conclusion that the measurements were taken in slightly different ways. In this paper I have used my own measurements, and, besides the above characters, the angles of the fundamental triangle, the indices involving B, $100\,G'H/GB$, $100\,NB/NH, R$, $100\,O_2/O_1R$ and $100\,G_2/G_1$ are altered from their original values. In addition to the means in Table I the following are now available for the two Naqada groups (see pp. 18, 19 below):

	L'	H	OH	S	S_3	S_3'
A and *Q* Series ...	185·4 (87)	136·4 (71)	115·6 (98)	372·6 (90)	116·9 (85)	96·5 (84)
B, *T* and *R* Series	182·8 (20)	134·7 (23)	113·8 (26)	372·3 (20)	116·9 (23)	97·0 (23)
	Q'	$G'H$	GB	NH, R	PH	EOW
A and *Q* Series ...	302·9 (87)	69·7 (52)	95·9 (60)	50·2 (54)	20·1 (56)	96·1 (56)
B, *T* and *R* Series	307·9 (18)	69·6 (18)	95·1 (18)	50·5 (19)	19·9 (17)	94·4 (18)
	G_1	G_1'	G_2	O_1R	O_2R	$100\,O_2/O_1R$
A and *Q* Series ...	52·5 (49)	48·0 (49)	40·7 (48)	41·6 (57)	32·2 (57)	{77·4 (46)}
B, *T* and *R* Series	50·7 (17)	46·1 (17)	40·3 (19)	41·7 (19)	32·1 (19)	{77·0 (19)}
	$100\,NB/NH, R$	$100\,G_2/G_1$	$100\,G_2/G_1'$	$100\,G'H/GB$	$100\,B/L'$	$100\,H/L'$
A and *Q* Series ...	{49·8 (54)}	{77·5 (48)}	{84·8 (48)}	{72·7 (52)}	{71·6 (87)}	{73·6 (71)}
B, *T* and *R* Series	{49·7 (18)}	{79·5 (17)}	{87·4 (17)}	{73·2 (18)}	{74·5 (20)}	{73·7 (20)}
	$100\,H/L$	$100\,B/H$	$P\angle$	θ_1	θ_2	*Oc. I.*
A and *Q* Series ...	{73·8 (71)}	{97·3 (71)}	84°·6 (50)	{29°·9 (46)}	{11°·3 (46)}	60·2 (84)
B, *T* and *R* Series	{73·2 (23)}	{101·0 (22)}	86°·1 (16)	{29°·6 (16)}	{12°·2 (16)}	60·5 (23)

* Wrongly given by Fouquet as 184·6. † Wrongly given by Fouquet as 88·1.

‡ ,, ,, ,, 37·7. § ,, ,, ,, 86·8.

‖ The means in this column are given by Chantre (15) in whole mms. only and it is probable that they were not calculated further, so they are not necessarily correct to the nearest mm. The mean indices were calculated from these curtailed lengths and they may easily differ from the true means by as much as one unit. There were 24 ♂ skulls in the series but some lengths, and in particular H' and J, could almost certainly not be taken on all of them. Chantre does not say how many skulls each mean was based on. The material is presented in such a thoroughly unsatisfactory way that it is practically worthless.

TABLE II.

Coefficients of Racial Likeness between Fouquet's Pre-dynastic Series.*

		El-Amrah (7·7)	Kawamil (14·2)	Beit-Allam (8·3)	Naqada South (25·9)
El-Amrah (7·7) ...	19 Characters ...	—	0·34 ± ·11	1·17 ± ·11	1·98 ± ·11
	6 Indices ...	—	0·16 ± ·19	1·65 ± ·19	2·79 ± ·19
Kawamil (14·2) ...	19 Characters ...	0·34 ± ·11	—	0·76 ± ·11	0·48 ± ·11
	6 Indices ...	0·16 ± ·19	—	1·65 ± ·19	0·39 ± ·19
Beit-Allam (8·3) ...	19 Characters ...	1·17 ± ·11	0·76 ± ·11	—	1·16 ± ·11
	6 Indices ...	1·65 ± ·19	1·65 ± ·19	—	2·65 ± ·19
Naqada South (25·9)	19 Characters ...	1·98 ± ·11	0·48 ± ·11	1·16 ± ·11	—
	6 Indices ...	2·79 ± ·19	0·39 ± ·19	2·65 ± ·19	—

* The numbers in brackets following the designation of the races are the mean numbers of crania available for the characters used in computing the Coefficients of Racial Likeness.

Although the series are short, the lowness of the Coefficients in Table II may be taken to indicate very close affinity. The Kawamil series most closely represents the mean type and the El-Amrah skulls stand rather apart from the others. Ten characters—viz. $100\,H'/L$, L, H', LB, U, J, NB, O_1', O_2 and fmb—of the 19 compared show no significant differences between the four series. The values of α† for the other nine characters are:

		$100\,B/L$	$100\,B/H'$	$100\,NB/NH'$	$100\,O_2/O_1'$	$100\,fmb/fml$	B	B'	NH'	fml
El-Amrah and	Kawamil	0·07	0·23	3·17	2·07	1·46	0·45	2·72	0·47	4·04
	Beit-Allam	3·66	2·47	1·86	1·32	5·59	2·71	6·86	1·77	10·29
	Naqada South	0·04	0·01	9·86	11·53	1·24	0·03	9·17	7·23	1·22
Kawamil and	Beit-Allam	6·62	5·09	0·08	0·30	2·36	7·34	1·70	0·68	2·76
	Naqada South	0·02	0·54	2·38	5·17	0·07	1·24	2·76	6·25	1·83
Beit-Allam and	Naqada South	7·17	3·47	2·48	4·44	3·38	2·60	0·01	1·57	8·20

The orbital is the only index which shows more than one value of α greater than five and the low value for the Naqada South series is responsible for the divergences in that case. The small frontal breadth (B') of the El-Amrah, the small nasal height (NH') of the Naqada South and the large foraminal length (fml) of the Beit-Allam series are also more or less distinctive. But by far the greater number of the cranial characters, and in particular those of the calvaria,

† With the usual notation $\alpha = \frac{n_s n_s'}{n_s + n_s'} \left\{ \frac{M_s - M_s'}{\sigma_s} \right\}$ and its value is taken to indicate a significant difference when it is greater than 5.

are so similar for the four series that they might represent identical types. It is reasonable to suppose that three of the series—the Kawamil, Beit-Allam and Naqada South—are quite pure representatives of the same race: their pooled means are given in Column 6 of Table I (p. 14). The El-Amrah skulls diverge slightly from that type. The standard deviations of the three pooled distributions may be compared with those of the long series of 26th to 30th Dynasties skulls which is dealt with in a later section of this paper.

	L	B	$100\,B/L$	B'	$100\,NB/NH'$
Fouquet's Pre-dynastic Series	5·79 ± ·36	5·06 ± ·31	3·11 ± ·19	6·41 ± ·39	4·69 ± ·35
26th to 30th Dynasty Series .	5·72 ± ·09	4·76 ± ·08	2·68 ± ·04	4·05 ± ·06	3·82 ± ·06*

* This S.D. is not of $100\,NB/NH'$ but of the very similar index $100\,NB/NH, R$.

The population made up by combining Fouquet's three series is certainly less homogeneous than that of the 26th to 30th Dynasties, but a comparison of the individual distributions suggests that the heterogeneity is inherent in each of them and is not due to the pooling. Provisionally we may assume that the pooled means are those of a pure racial type and we may speak of that as the Egyptian Aeneolithic type*. It will be shown that this primitive Aeneolithic race is most clearly differentiated from all the racial types that inhabited Egypt in Dynastic times and it must have been exterminated altogether, or, as the cultural evidence suggests, absorbed in a much larger population. The type is dolichocephalic and among African races it is distinguished by its large calvaria. Many modern races equal it in skull length, but apparently none but primitive types have as great a skull base (nasio-basial length) ($LB = 105{\cdot}5$). The Eskimo LB length is 104·9, the Aino 105·4 and the Moriori 105·6†. It is reasonable to suppose that this primitive Egyptian Aeneolithic race was directly descended from the people whose palaeoliths are found on the high terraces on either side of the Nile Valley.

At least one other race was living in Upper Egypt at the time when the Aeneolithic people were still distinct in type. In the well-known work of Thomson and MacIver (13) individual measurements of 1500 skulls are provided. The material is divided up into series dating from Early Pre-dynastic to Roman times. Unfortunately the measurements are few in number and many of them are not defined precisely although there can be no doubt that they correspond exactly to measurements that are ordinarily taken in the Biometric Laboratory.

The Pre-dynastic skulls are divided into two groups, an early and a late; the relative ages having been determined by the method of Petrie's sequence dates. Of the ♂ skulls, 20 of the Early Pre-dynastic came from Abydos, 10 from El-Amrah and 25 from Hou: the pooled means are shown in Column 9 of Table I. Twenty-

* Although a few small metal objects were found in the early graves opened by J. de Morgan, he supposed that they belonged to a Neolithic stage of culture.

† These values are taken from *Biometrika*, Vol. XIV. p. 222.

eight of the Late Pre-dynastic ♂ specimens came from El-Amrah and 99 from Hou; their pooled means are in Column 10*.

Several large Early Pre-dynastic cemeteries were excavated by Petrie and Quibell in 1895 and an account of the work was given in the following year†(3). Fawcett's paper of 1901 (6) dealt with 300 of these skulls. They had been taken from at least four distinct cemeteries within a radius of three or four miles, but the grave furniture found at the different sites was so similar that the fact that a homogeneous population was being dealt with appeared to be conclusively established. An examination of the skull measurements led to the same conclusion. The standard deviations of characters are quite as low as for populations which craniologists accept as homogeneous, though the majority of them are slightly greater than those of a long series—the *E* series—of Egyptian skulls of the 26th to 30th Dynasties which will be dealt with in a later section of this paper‡. A comparison of the mean measurements of the series from different cemeteries suggested, however, that the divergences, though small, might be of significance when considered in relation to the quantity of material that has become available since Fawcett wrote. By far the greater number of the so-called Naqada crania came from two cemeteries: one near the village of Nubt and the other two miles to the north towards Ballas§. These form the first series—the skull numbers not preceded by a letter and which for convenience we will term the *A* series—and the *Q* series respectively in Fawcett's tables of individual measurements. The *T* and *B* series were from two distinct burial groups within half a mile of the first site. The mean measurements of the *A* and *Q* series are almost identical and their combined means‖ are shown in Column 7 of Table I. The *T*, *B* and *R* series are very similar to one another and they agree with the *A* and *Q* series for the majority of characters but are yet sufficiently remote for others to make a further enquiry desirable. The splitting up of a relatively short series of crania into groups which are supposed more or less racially distinct, when the whole material came from a single district and was homogeneous from the standpoint of the archaeological evidence, is a practice that the biometric craniologist is bound to discredit unless very strong evidence can be adduced to support it. There are certainly not two *distinct* types

* In the earlier work of MacIver (12) measurements of smaller numbers of Pre-dynastic crania were provided. The means of these—taking Early and Late periods together—were reduced by Miss M. A. Lewenz and Professor Karl Pearson and quoted by Fawcett ((6), p. 426). Individual measurements of 33 of the 44 skulls were given again in (13), so the earlier series is no longer of any importance; it is obviously intermediate between our Early and Late Pre-dynastic types. MacIver gave orbital measurements that were not reproduced in (13), but when divided up the numbers become so small that I have omitted to consider them.

† It was orignally thought that the Naqada people were of Early Dynastic date but further evidence showed that they must be assigned to the Pre-dynastic period.

‡ The standard deviations of the Naqada, the *E* and three other series are compared in *Biometrika*, Vol. xiv. p. 208.

§ The "New Race" cemeteries excavated by Petrie and Quibell were four or five miles north of the village of Naqada. The sites are shown in (3) Plates I and II. J. de Morgan worked in Naqada itself and it was there that he found Menes and the Naqada South skulls.

‖ With the modifications of Fawcett's measurements explained in the notes to the Table itself.

among the Naqada crania. There were two distinct racial types in early Egypt, but they were quite as similar to one another as adjacent racial populations are generally found to be. Nearly all our cranial series represent intermediate types between those two and the differences between them are very small; so small that with less abundant material it would have been impossible to say that they had a real statistical significance. Our contention is that Fawcett's Naqada series of skulls contains representatives of two slightly differing populations which are both intermediate between the two extreme Egyptian types, and the evidence in favour of that view is partly deduced from a comparison of the two groups themselves, but its real justification is found in comparing their affinities with other early series. The Coefficients between the Naqada groups and Thomson and MacIver's Pre-dynastic series are given in the following Table:

TABLE III.

Coefficients of Racial Likeness between some Pre-dynastic Series.

		Early Pre-dynastic (40·9)*	Naqada *A* and *Q* Series (66·1)	Late Pre-dynastic (105·6)	Naqada *B*, *T* and *R* Series (18·6)
Early Pre-dynastic (40·9)	14 Characters ...	—	0·31 ± ·13	2·17 ± ·13	3·55 ± ·13
	6 Indices and Angles	—	0·53 ± ·19	2·09 ± ·19	6·04 ± ·19
Naqada *A* and *Q* Series (66·1)	14 Characters ...	0·31 ± ·13	—	0·73 ± ·13	1·62 ± ·13†
	6 Indices and Angles	0·53 ± ·19	—	1·62 ± ·19	3·25 ± ·19
Late Pre-dynastic (105·6)	14 Characters ...	2·17 ± ·13	0·73 ± ·13	—	1·66 ± ·13
	6 Indices and Angles	2·09 ± ·19	1·62 ± ·19	—	3·40 ± ·19
Naqada *B*, *T* and *R* Series (18·6)	14 Characters ...	3 55 ± ·13	1·62 ± ·13†	1·66 ± ·13	—
	6 Indices and Angles	6·04 ± ·19	3·25 ± ·19	3·40 ± ·19	—

* See footnote to Table II, p. 16.

† These Coefficients are based on 14 characters and 6 indices and angles so that they may be comparable with the others in the table. When they are calculated for the full 31 characters, of which 12 are indices and angles, they become 0·65 ± ·09 and 1·54 ± ·14, so they are substantially reduced. The added characters show no significant differences.

The low order of the Coefficients in Table III shows that we are dealing with closely related types. In comparing them it is important to pay due regard to the number of skulls available for each series—the number in brackets following the designation of the race. The Coefficients involving the Naqada *B*, *T* and *R* series are all symptomatic of greater racial divergences than those between any pair of the other three types, while the former stands closest to the Late Pre-dynastic. The Naqada *A* and *Q* group is almost identically the same as the Early Pre-dynastic, but between the latter and the Naqada *B*, *T* and *R* group there is a difference in type on which considerable stress may be laid as the numbers of skulls in the two

series are small. The Coefficients arrange the series in the order shown in the table and that order is a perfectly consistent one. Turning to a comparison of the individual characters we find that 8 of the 14 compared—viz. $100\,NB/NH'$, L, H', LB, $G'H$, NH', NB and $A\angle$—show no significant differences whatever. The only values of α greater than 5 are for the characters shown below.

		B	$100\,B/L$	$100\,B/H'$	J	$100\,H'/L$	$N\angle$
Early Pre-dynastic and	Naqada A and Q ...	2·45	0·07	1·63	3·93	2·39	4·14
	Late Pre-dynastic...	6·55	0·82	9·06	12·74	7·53	0·11
	Naqada B, T and R	15·13	11·54	18·79	1·80	2·82	7·37
Naqada A and Q and	Late Pre-dynastic ...	1·14	0·62	3·85	1·51	1·76	4·81
	Naqada B, T and R...	8·96	11·95	11·18	0·04	0·42	1·56
Late Pre-dynastic and	Naqada B, T and R	6·00	9·47	5·21	0·40	0·02	7·83

The most significant differences are between the calvarial and zygomatic breadths and the indices $100\,B/L$ and $100\,B/H'$ which vary with B while H' and L remain constant. No stress can be laid on the single significant difference between the indices $100\,H'/L$. The two values of α greater than 5 in the case of the nasal angle are of greater significance. The actual measurements are:

	B	$100\,B/L$	$100\,B/H'$	J	H'	$N\angle$
Early Pre-dynastic	131·4	71·7	98·1	124·3	134·0	67°·0
Naqada A and Q...	132·7	71·8	99·2	125·9	133·8	65°·5
Late Pre-dynastic	133·4	72·1	100·5	127·0	132·7	66°·8
Naqada B, T and R	136·1	74·0	103·0	126·2	132·2	64°·3

The gradation shown in the case of the first three of these characters confirms the reasonableness of the order suggested by the Coefficients. H', though showing no significant differences, exhibits the same order. The discordance in the sequence of zygomatic breadths may well be attributed to the errors of random sampling. The variability of the nasal angles would appear to be due to some other cause and it is almost certainly the result of slight differences between the methods of measuring the sides of the fundamental triangle employed by the workers concerned*. The differences between the two Naqada groups are precisely similar to those which distinguish them from Thomson and MacIver's two series. From Early to Late Pre-dynastic times—a period of at least 2000 years—the change in the skull form of the population of Upper Egypt was exceedingly small. It may have been due to slow, progressive evolution or to slight admixture with another strain. A comparison of these Pre-dynastic types with Fouquet's Aeneolithic will be of interest. The Coefficients are:

* Probably due to differences in the determination of the alveolar point: see Note p. 53.

TABLE IV.

Coefficients of Racial Likeness between Fouquet's Aeneolithic and other Pre-dynastic Series.

		Early Pre-dynastic (40·9)	Naqada *A* and *Q* (66·1)	Late Pre-dynastic (105·6)	Naqada *B*, *T* and *R* (18·6)
Aeneolithic (48·7)	All Characters ...	11·23 ± ·14	9·18 ± ·11	10·71 ± ·14	5·28 ± ·11
	Number of Characters	11	19	11	19
	Indices and Angles ...	1·31 ± ·24	0·29 ± ·19	1·66 ± ·24	1·41 ± ·19
	Number of Characters	4	6	4	6

The most striking feature of Table IV is the great difference between the values of the Coefficients for all characters and for indices and angles only. With Thomson and MacIver's and Fouquet's series the only four indices that can be compared are $100\,B/L$, $100\,H'/L$, $100\,B/H'$ and $100\,NB/NH'$, while two others—$100\,O_2/O_1'$ and $100\,fmb/fml$—take part in the comparison with the Naqada groups. Only one difference of indices is significant; that is with the Naqada *B*, *T* and *R* series where $\alpha = 6{\cdot}87$ for $100\,B/H'$. The high values of the Coefficients for all characters are due entirely to the greater size of the Aeneolithic calvaria: facial direct measurements show no significant differences whatever. So the Aeneolithic type was almost identical with the Pre-dynastic Egyptian in size and shape of face and shape of calvaria, but it greatly exceeded the more advanced type in absolute size of calvaria*. The relation is a curious one and we can offer no explanation of it, but must consider the types to be perfectly distinct. The Aeneolithic people may have been absorbed into the larger population, but there is not the slightest suggestion that the latter was thereby modified in any way.

In the last column of Table I the mean measurements of a series of Pre-dynastic skulls discovered at El-Khozan and measured by E. Chantre are quoted from (15), p. 57. The measurements are presented in such an altogether unsatisfactory way that it is not possible to determine the relationships of the type with any certainty†. From a study of the grave finds Chantre supposed that the skulls belonged to the same race and were contemporary with J. de Morgan's series (see (14)) and he said that this conclusion was confirmed by the similarity of the skulls themselves. But the mean measurements are obviously very dissimilar to those of Fouquet's Aeneolithic type. They apparently resemble the Late Pre-dynastic type more than any other. No stress can be laid on the significance of the small H' and J of the El-Khozan skulls as either mean may easily have been based on no more than half of the 24 specimens.

* Judging from the drawings provided by Fouquet in appendices to J. de Morgan's volumes (10) and (11) the skulls of the Aeneolithic type had well-marked superciliary ridges, mastoids and inionic prominences and were decidedly more muscular than dynastic Egyptian crania.

† See notes to Table I.

A comparison of all the available skull measurements leads to the conclusion that in Pre-dynastic times there were two distinct racial types in Upper Egypt. The first and more primitive was characterised by a large calvaria and it was probably represented by a small and dwindling population. The second, which was carried down to historic times, was being transformed by some slow process of evolution or racial admixture. Both main types had low cephalic indices which distinguish them from the type of all later Egyptians. The apparent relationships of the various series are represented schematically in Fig. 1, p. 2.

4. *Skulls of the First to the Sixth Dynasties.* The first six dynasties cover the protodynastic era and the Old Kingdom. At the beginning of that period a change in the administration of the country took place which was likely to be of considerable ethnic importance. The kingdoms of Upper and Lower Egypt had been ruled by one king before the accession of Menes, the founder of the First Dynasty, but he unified the country and facilitated the intercourse between the north and the south.

All the available mean measurements of Early Dynastic series of skulls are shown in Table V. The series in the third and later columns is one which was measured by Mr G. H. Motley and the means are now given for the first time*. The collection consists of 48 skulls excavated by Sir Flinders Petrie in 1922 and presented by him to the Biometric Laboratory. They belonged to court-officials of the reigns of Zev and Zet, third and fourth kings of the First Dynasty. The bodies were buried in a square, the graves forming the border or sides, at Abydos between the Shuneh and the Coptic Deir. Each king had a fresh square of burials†. The full account of these excavations will be given in a volume entitled *Tombs of the Courtiers*, to be published shortly by Sir Flinders Petrie.

Of the 48 skulls one (No. 443) was almost certainly that of a negro; it is possible that some of the others showed, but far less definitely, negroid characters. Of the remaining 47 crania, 37 were allotted to the male, 10 to the female sex. Actually 81 crania were obtained and in sufficiently good condition to be treated with paraffin preparatory to transport; through some misadventure only 48 were packed and reached England. A series of 15 measurements were taken by Sir Flinders Petrie in the field; seven of these, our L (the maximum length), B (maximum breadth), H' (the basio-bregmatic height), LB (the basio-nasal height), $G'H$ (upper face height, nasion to alveolar point), GL (profile length from basion to alveolar point) and J (the bizygomatic breadth), are customary cranial measurements. The remaining eight measurements were taken for comparison with statuary, and it is less certain that they are: *Orbital Breadth* (? our O_1), *Jaw Breadth* (? our condylar width W_1), *Facial Height* (Nasion to base of jaw), *Jaw Height* (gnathion to infradental + height of teeth), *Nasidental Height* (nasion to line of teeth), *Nasal Height* (? our NH'), *Internal Orbital Width* and *External Orbital Width*, both measured

* Measurements of the individual skulls and remarks on cranial anomalies will be found in the Appendix to this paper.

† There is reason to believe that at least some of the interments must, by the attitude of the skeletons, have been cases of persons buried alive. An entire skeleton, to illustrate this, was brought back by Sir Flinders Petrie and is now mounted in the Museum of the Laboratory.

inside the orbits. As we invariably measure the external and internal ocular distances on the living subject, it seems highly important to include these among craniological measurements. I am inclined to think that the external orbital distance is better not taken *inside* the orbit, as then for comparison with the living we must allow for thickness of flesh; it would, I suggest, be advisable to take it from the inside border of one orbital ridge to that of the other, the measurement to be taken at the fronto-malar sutures, which can be closely approximated to in the living. The terminals of the internal orbital measurement are less easy to standardise. We have already two craniological measurements of this kind, namely *DC* the dacryal chord, and *SC* the minimum chord between the two naso-maxillary sutures, the so-called simotic chord. There is no doubt that the dacryal chord would better represent the internal orbital distance, and might be adopted for this purpose to save multiplication of cranial measurements*. If we add to the dacryal chord the widths of the two orbits we have for the mean total $\begin{cases} ♀ 21{\cdot}56 + 40{\cdot}65 + 40{\cdot}09 = 102{\cdot}30 \\ ♂ 22{\cdot}82 + 42{\cdot}77 + 42{\cdot}33 = 107{\cdot}92 \end{cases}$. But this is not a true measurement of the external ocular distance as the three segments are not in the same straight line. Using the measurement indicated above we find that the external ocular distance for males = 98·4 and for females 93·2 mms., while Sir Flinders Petrie's measurements for these 37 and 10 crania are 96·4 and 92·1 mm., so that taking the measurements "inside the orbits" makes a difference of 2 mm. in the male and 1 mm. in the female. The measurement from one fronto-malar to the other fronto-malar point seems to be somewhat in excess of the "inside the orbits" external ocular distance. It has the advantage, however, of being both an easily determined length on the skull and also a readily measured quantity on the living head. On the crania we have compared, Sir Flinders Petrie's field measurements, when taken for the same characters as those used by us, agree reasonably, 0·5 to 1 mm. being the usual difference†. A much greater divergence occurred in the sexing. We were at first much inclined to make the whole series *male*, but have finally admitted ten to be female. There is no archaeological evidence of sex, and failing this, we feel compelled to retain as a minimum four-fifths of the material as male. It is of course difficult with a short series of skulls, where we are not compelled by the circumstances of the interment—very peculiar in this case—to suppose both sexes represented, or represented in approximately equal numbers, to draw a marked line anatomically across the series separating it into two sex groups. It was indeed not hard to classify the material into two categories, and these might have been regarded as a sex-division; but they appeared to us to correspond far more to racial than to sexual differences. In other words we doubt the homogeneity of our material, we believe it to consist of a mixture of races, or better a race with individuals of more or less mixed blood.

* Our value of the dacryal chord for 34 ♂'s is 22·8 mm. Sir Flinders Petrie's internal orbital width, measured "inside the orbits," is 25·6 mm. for the same crania. This would only give a difference of 0·7 mm. in determining the distance of the pupillary centre from the sagittal plane and would serve for portraiture.

† The field measurements were taken before paraffining, but that process seems to have made no appreciable difference to the skulls.

TABLE V.

Mean Measurements of Egyptian Series from the First to the Fifth Dynasties. Males.*

Dynasty ...	1st Private Tombs	1st Royal Tombs	1st and 2nd Royal Tombs	4th and 5th	4th	Dynasty	1st Royal Tombs	Dynasty	1st Royal Tombs
Locality ...	Abydos	Abydos	Abydos	Deshasheh and Medum	Sakkarah	Locality	Abydos	Locality	Abydos
Measured by...	Thomson and MacIver	Motley	Thomson and MacIver	Thomson and MacIver	Broca	Measured by	Motley	Measured by	Motley
F	182·4 (43)	182·3 (36)	182·2 (27)	183·0 (30)	—	Capacity	1364·3 (24)	G_1'	49·5 (32)
L	183·8 (43)	184·9 (36)	184·8 (27)	184·9 (54)	185 (31)	*W* (grs.)	668·7 (32)	G_2	42·7 (29)
B	133·7 (41)	136·2 (37)	137·7 (26)	139·3 (54)	141 (31)	*L'*	185·5 (31)	*EB*	42·8 (29)
B'	—	93·0 (35)	—	—	95 (31)	*H*	133·8 (30)	*EH*	14·7 (29)
H'	135·1 (38)	132·1 (35)	132·8 (24)	136·0 (50)	137 (31)	*OH*	110·3 (31)	100 *B*/*L'*	73·4 (31)
LB	102·3 (30)	102·0 (33)	100·1 (19)	101·8 (47)	—	*Q*	304·9 (31)	100 *H*/*L'*	72·1 (30)
U	—	513·5 (30)	—	—	518 (31)	*Q'*	305·5 (30)	100 *H*/*L*	72·3 (30)
G'H	70·6 (29)	74·7 (31)	72·2 (20)	71·3 (19)	—	*S*	369·0 (29)	100 *B*/*H*	102·1 (30)
J	130·9 (25)	129·7 (31)	129·9 (16)	127·1 (16)	131 (31)	S_1	123·7 (32)	100 *G'H*/*GB*	76·8 (30)
NH'	51·4 (30)	53·7 (32)	51·9 (21)	52·0 (44)	52 (31)	S_2	129·4 (34)	100 *NB*/*NH*, *R*	47·5 (31)
NB	25·8 (30)	25·5 (33)	25·2 (21)	25·1 (43)	24 (31)	S_3	116·9 (32)	100 *NB*/*NH*, *L*	47·7 (32)
DC	—	22·8 (34)	—	—	24 (31)	*PH*	21·3 (34)	100 O_2/O_1, *R*	78·2 (33)
$O_1'R$	—	40·0 (32)	—	—	38 (31)†	*GB*	97·3 (33)	100 O_2/O_1, *L*	80·4 (31)
O_2R	—	33·5 (33)	—	—	33 (31)†	*NH*, *R*	53·9 (32)	100 *DS*/*DC*	32·6 (29)
GL	98·6 (29)	96·2 (30)	96·5 (19)	96·3 (42)	—	*NH*, *L*	54·0 (32)	100 *SS*/*SC*	42·4 (30)
fml	—	37·0 (33)	—	—	36 (31)	*DS*	12·2 (29)	100 G_2/G_1	78·6 (29)
fmb	—	30·0 (32)	—	—	30 (31)	*DA*	35·8 (29)	100 G_2/G_1'	{86·4 (29)}
100 *B*/*L*	{72·7 (41)}	73·8 (36)	{74·9 (26)}	{75·4 (50)}	{76·2 (31)}	*SS*	4·58 (30)	100 *EH*/*EB*	34·4 (29)
100 *H'*/*L*	{73·6 (38)}	71·3 (34)	{72·0 (23)}	{72·6 (50)}	{74·1 (31)}	*SC*	10·97 (34)	*P* ∠	86°·1 (32)
100 *B*/*H'*	{99·0 (38)}	103·6 (35)	{103·7 (24)}	{102·4 (50)}	{102·9 (31)}	O_1R	42·8 (33)	θ_1	30°·3 (30)
100 (*B*−*H'*)/*L*	{−0·9 (38)}	2·64 (34)	{2·9 (23)}	{1·8 (50)}	{2·2 (31)}	O_1L	42·3 (31)	θ_2	13°·9 (30)
100 *NB*/*NH'*	{50·4 (30)}	47·9 (32)	{48·5 (20)}	{48·2 (44)}	{46·1 (31)}	O_2R	33·5 (33)	S_3'	96·0 (32)
100 O_2/O_1', *R*	—	86·1 (32)	—	—	{86·8 (31)}†	O_2L	33·9 (31)	*Oc. I.*	59·9 (32)
100 *fmb*/*fml*	—	81·5 (32)	—	—	{83·3 (31)}	*EOW*	98·4 (37)		
N ∠	{66°·2 (29)}	63°·4 (30)	{65°·5 (19)}	{64°·6 (19)}	—	G_1	54·6 (32)		
A ∠	{72°·8 (29)}	72°·4 (30)	{71°·5 (19)}	{73°·4 (19)}	—				
B ∠	{41°·0 (29)}	44°·2 (30)	{43°·0 (19)}	{42°·0 (19)}	—				

* The angles and indices in curled brackets were found from the mean measurements of the component lengths instead of from the individual measurements.

† These measurements may be of either the right or left orbit.

Without venturing to be in the least dogmatic, on such slender evidence, we should expect a divergence from the prehistoric Egyptians by their mixture with a race of a coarser, stronger type. If the line drawn across the material corresponds to more and less of this stronger coarser race, it will be obvious that there is danger in such a small series of confusing such a racial line with a sex line. Were anyone to assert on other grounds than we have had before us that the whole series is male, there are certainly only a very few crania about which we should be inclined to contest his assertion.

Measurements of two other series of First Dynasty skulls have been furnished by Thomson and MacIver (13). One came from a private graveyard and the other from the Royal Tombs of the First and Second Dynasties; both were from Abydos. The same writers have provided measurements of Fourth and Fifth Dynasties skulls from Deshasheh and Medum in Lower Egypt. The mean measurements are shown in Column 5 of Table V.

The Coefficients of Racial Likeness between the above series were calculated. The numbers of skulls are fairly adequate but unfortunately there are no more than 14 characters, of which 6 are indices and angles, that can be used for that purpose.

TABLE VI.

Coefficients of Racial Likeness between Early Dynastic Series.

		1st Dynasty Private Tombs (33·6)	1st Dynasty Royal Tombs (Motley) (33·2)	1st and 2nd Dynasties Royal Tombs (21·8)	4th and 5th Dynasties (39·9)
1st Dynasty ... Private Tombs (33·6)	14 Characters ...	—	5·37 ± ·13	3·54 ± ·13	5·59 ± ·13
	6 Indices and Angles	—	7·72 ± ·19	5·28 ± ·19	6·97 ± ·19
1st Dynasty ... Royal Tombs (33·2) ...	14 Characters ...	5·37 ± ·13	—	0·66 ± ·13	2·80 ± ·13
	6 Indices and Angles	7·72 ± ·19	—	0·36 ± ·19	1·62 ± ·19
1st and 2nd Dynasties Royal Tombs (21·8) ...	14 Characters ...	3·54 ± ·13	0·66 ± ·13	—	0·39 ± ·13
	6 Indices and Angles	5·28 ± ·19	0·36 ± ·19	—	0·07 ± ·19
4th and 5th Dynasties (39·9)	14 Characters ...	5·59 ± ·13	2·80 ± ·13	0·39 ± ·13	—
	6 Indices and Angles	6·97 ± ·19	1·62 ± ·19	0·07 ± ·19	—

The most striking feature of this table is the highness of the Coefficients between the First Dynasty skulls from the private tombs and the other series. The numbers of skulls being small, the Coefficients are of an order which must be taken to indicate perfectly distinct racial types. The two sets of skulls from the royal cemeteries are very similar to one another and are far more closely related to the later type from Lower Egypt than to the contemporary civil population of their own town. The Coefficient between the First and Second Dynasties series and the Fourth and Fifth Dynasties series is as low as it might have been if the two had

been drawn from the same population. Motley's series, however, is distinct from that of the Fourth and Fifth Dynasties. From this evidence alone, which by itself would be very scanty of course, we may suppose that the courtier population of Abydos in the First and Second Dynasties was intrusive and that it came from Lower Egypt where the type persisted until at least as late as the Fifth Dynasty. A comparison of the individual characters of the four series can be made in the usual way by considering the values of the α's between them. There are only four characters—viz. *L*, *LB*, *NB* and $A\angle$—which show no values of α greater than 5. The α's for the remaining ten are given in the table below.

All the greatest values of α are in the first three rows of this table. The indices $100\,B/H'$ and $100\,B/L$ of the First Dynasty Private Tombs series are quite distinct from the values for the other three types and the divergences are evidently due primarily to differences in calvarial breadth. In the last rows there are only five characters with values of α greater than 5. In the whole table the only significant differences of $G'H$ and NH' are those from Motley's First Dynasty series. The mean facial and nasal heights of those skulls are greater, and in the majority of cases, quite significantly greater than those of any other ancient Egyptian series; and other characters, not provided by Thomson and MacIver, also distinguish them, so there is clear evidence of admixture with some alien strain. At the same time they stand very close to the other Royal Tombs series of the First and Second Dynasties and they are primarily of Egyptian type. The Fourth and Fifth Dynasties series from Lower Egypt is distinguished from the royal retainers of the First and Second Dynasties by having a much greater basio-bregmatic height.

		$100\,B/H'$	$100\,B/L$	B	$100\,H'/L$	$100\,NB/NH'$	H'	$G'H$	NH'	J	$N\angle$
1st Dynasty	1st Dynasty Royal Tombs (Motley)	20·84	3·26	5·35	10·98	6·51	5·01	14·64	9·60	0·95	10·54
Private Tombs	1st and 2nd Dynasties Royal Tombs	17·56	10·80	11·22	4·24	2·92	2·38	1·76	0·36	4·68	0·52
	4th and 5th Dynasties	13·49	23·02	32·18	2·50	5·80	0·51	0·32	0·75	6·74	2·69
1st Dynasty Royal Tombs (Motley)	1st and 2nd Dynasties Royal Tombs	0·01	2·60	1·51	0·78	0·31	0·21	4·41	4·82	0·02	4·68
	4th and 5th Dynasties	1·61	7·51	9·29	3·97	0·11	9·57	7·92	6·28	3·41	1·52
1st and 2nd Dynasties Royal Tombs	4th and 5th Dynasties	1·48	0·60	1·98	0·66	0·08	5·08	0·46	0·01	3·00	0·70

The Coefficients between these Early Dynastic and the four Pre-dynastic series previously considered are given in Table VII.

TABLE VII.

Coefficients of Racial Likeness between Early Dynastic and Pre-dynastic Series.

		1st Dynasty Private Tombs (33·6)	1st Dynasty Royal Tombs Motley (33·2)	1st and 2nd Dynasties Royal Tombs (21·8)	4th and 5th Dynasties (39·9)
Early Pre-dynastic (40·9)	14 Characters ...	2·50 ± ·13	12·15 ± ·13	7·64 ± ·13	11·86 ± ·13
	6 Indices and Angles	−0·11 ± ·19	12·17 ± ·19	8·62 ± ·19	13·36 ± ·19
Naqada *A* and *Q* (66·1)	14 Characters ...	1·92 ± ·13	9·68 ± ·13*	5·98 ± ·13	10·79 ± ·13
	6 Indices and Angles	0·46 ± ·19	8·60 ± ·19	7·81 ± ·19	12·45 ± ·19
Late Pre-dynastic (105·6)	14 Characters ...	2·17 ± ·13	9·82 ± ·13	4·88 ± ·13	11·42 ± ·13
	6 Indices and Angles	1·84 ± ·19	9·99 ± ·19	6·55 ± ·19	13·75 ± ·19
Naqada *B*, *T* and *R* (18·6)	14 Characters ...	2·45 ± ·13	2·59 ± ·13*	0·56 ± ·13	1·27 ± ·13
	6 Indices and Angles	2·78 ± ·19	0·22 ± ·19	0·37 ± ·19	0·33 ± ·19

* The Coefficients between Motley's First Dynasty series and the two Naqada can also be calculated for the whole 31 characters, of which 12 are indices and angles; with the *A* and *Q* they are 6·36 ± ·09 and 6·65 ± ·14 and with the *B*, *T* and *R* 2·37 ± ·09 and 0·90 ± ·14 respectively for all characters and for indices and angles only.

The Coefficients in Table VII show that there is a distinct line of cleavage dividing the series into two opposed groups. The Private Tombs series of the First Dynasty belongs clearly to the Pre-dynastic stock and it is apparently most closely related to the Late Pre-dynastic type which is represented by more skulls than the other two. And those four series are most clearly differentiated from the other three arranged horizontally in the table; the latter form another group of closely inter-related types. For convenience we may term these groups the Upper—the Pre-dynastic—and Lower Egyptian types respectively; the appropriateness of the titles will be more apparent when later series are considered. The Naqada *A* and *Q* group behaves in every way like Thomson and MacIver's Pre-dynastic types but the distinctiveness of the *B*, *T* and *R* skulls is again brought out in Table VII. The series being a short one its Coefficients are for the most part of a lower order than the others in the table but it clearly inclines more to the Lower than to the Upper Egyptian type and may be considered intermediate between the two. The division between the two main groups is quite sharp, but there is very definite evidence to show that the transformation from one to the other had started in prehistoric times. A comparison of all the Coefficients shows that the Early Pre-dynastic is the series most widely removed from the Lower Egyptian type and the Naqada *A* and *Q* group is more distant than the Late Pre-dynastic. A similar gradation, providing the principal justification for the arrangement shown in Fig. 1, p. 2, is found among the Lower Egyptian series. The relationships are brought out more clearly by comparing individual characters. Whether comparison is made between

series within either group or between series of different groups, the same characters always show the most significant differences and the same are constant throughout so long as Motley's First Dynasty series is not concerned in the comparison. From the latter, greatest divergences—i.e. the highest values of α—are in every case for the characters $G'H$ and NH', but otherwise it behaves like a pure Egyptian series*. In general the relationships between these series are such that:

(i) L, LB, NB and $A\angle$ show no significant differences whatever,

(ii) $100\,H'/L$, $G'H$ and NH' are practically constant,

(iii) $N\angle$, H' and $100\,NB/NH'$ have small but orderly differences, and

(iv) the greatest differences are found always for B, $100\,B/H'$, $100\,B/L$ and to a less marked extent for J.

The measurements showing most significant differences are compared below:

		B	$100\,B/L$	$100\,B/H'$	J	H'	$100\,NB/NH'$	$N\angle$
Upper Egyptian Type	Early Pre-dynastic	131·4	71·7	98·1	124·3	134·0	50·3	67°·0
	Naqada *A* and *Q*	132·7	71·8	99·2	125·9	133·8	50·0	65°·5
	Late Pre-dynastic	133·4	72·1	100·5	127·0	132·7	51·0	66°·8
	1st Dynasty Private Tombs	133·7	72·7	99·0	130·9†	135·1	50·4	66°·2
	Naqada *B*, *T* and *R*	136·1	74·0	103·0	126·2	132·2	49·9	64°·3
Lower Egyptian Type	1st Dynasty Royal Tombs (Motley)	136·2	73·8	103·6	129·7	132·1	47·9	63°·4
	1st and 2nd Dynasties Royal Tombs	137·7	74·9	103·7	129·9	132·8	48·5	65°·5
	4th Dynasty (Broca)	137	76·2	102·9	131	137	46·1	—
	4th and 5th Dynasties	139·3	75·4	102·5	127·1†	136·0	48·2	64°·6

† These bizygomatic breadths are based on fewer crania than the corresponding calvarial breadths. In the case of the First Dynasty series B is the mean of 41 skulls and J of 25. The mean B for those 25 is 134·8 so the J for the 41 would almost certainly have been greater than the recorded value and more in accordance with those of the allied series. In the same way it is probable that the Fourth and Fifth Dynasties J should be greater than 127·1.

The most important differences between the Upper and Lower Egyptian types are those associated with the breadth of the calvaria—B, $100\,B/L$, and $100\,B/H'$; the latter is the broader and less dolichocephalic skull. Those differences are correlated with lesser differences between the zygomatic breadths. The Lower Egyptian type also has the lesser nasal index and nasal angle. The basio-bregmatic

* We are only considering here the 14 characters measured by Thomson and MacIver which can be used in calculating the Coefficients of Racial Likeness. There are other characters which distinguish Motley's First Dynasty series from all purer Egyptian types (see below pp. 42 and 43).

heights are less regular. The values of H' for the Fourth Dynasty series are quite significantly greater than those for the First Dynasty Royal Tombs skulls and it will be shown later that the smaller values are those of the pure type while the others can only be accounted for on the supposition that the skulls represent a slight admixture with a race foreign to Egypt. So we may conclude that the Upper Egyptian type has a greater basio-bregmatic height than the Lower. The Fourth Dynasty series from Sakkarah, of which measurements are given above and in Column 6 of Table V, was measured by Broca. The measurements are in his unpublished note-books but the means have been quoted by several French craniologists*. Unfortunately they are given in whole mms. only and perhaps not to the nearest mm., while it is not known how many of the 31 skulls each mean is based on. The means are very similar to those of the contemporary series from Deshasheh and Medum measured by Thomson and MacIver†.

It will be convenient to consider first two later series from Lower Egypt. There are no others from that region: the majority of the ancient Egyptian skulls that have been measured came from the Thebaid.

5. *Series from Lower Egypt later than the Fifth Dynasty.* A long series of crania belonging to the period from the 26th to 30th Dynasties and containing upwards of 1700 individuals has been measured in the Biometric Laboratory by Miss A. Davin and others, and a paper on this series has recently been published‡. The object of that paper was to determine authoritative values for the correlation of cranial characters. This object can only be obtained by the measurement of a long series. The present use of that series is only a by-product of that enquiry, which has itself no direct bearing on the present paper. The standard deviations found for the E series have, however, been used in this and in other recent craniological work as constants for Coefficients of Racial Likeness. This is done because, being deduced from large numbers of crania, the values are more likely to express the general variation of the human skull within a fairly homogeneous group than standard deviations deduced from the small series which are all it is feasible to obtain as samples in the case of many racial groups. The mean measurements of the 26th to 30th Dynasties skulls are given in Table XIV. The specimens came

* The means given in Table V were taken from (16), p. 507 and (15), p. 57.

† In the winter of 1924—5 the British School excavated sites at Qau, 30 miles south of Assiut, of the Fourth to the Ninth Dynasties. The human remains discovered are of particular interest because there are no records of others from Middle Egypt. Through the courtesy of Sir Flinders Petrie I am able to give here the means of his field measurements of some of the ♂ skulls. The series has been presented to the Biometric Laboratory and a full report on it will appear later. The means are: $L = 184{\cdot}7$ (31), $B = 139{\cdot}6$ (31), $H' = 136{\cdot}9$ (30), $J = 127{\cdot}7$ (28), $DC = 24{\cdot}6$ (30), $GL = 95{\cdot}5$ (29), $LB = 101{\cdot}2$ (30), $G'H = 70{\cdot}3$ (29), $NH'? = 51{\cdot}2$ (29), $NB = 25{\cdot}4$ (29), $100\,B/L = \{75{\cdot}6\ (31)\}$, $100\,H'/L = \{74{\cdot}1\ (30)\}$, $100\,B/H' = \{102{\cdot}0\ (30)\}$, $100\,NB/NH' = 49{\cdot}6$ (29), $N\angle = \{64°{\cdot}5\ (29)\}$, $A\angle = \{73°{\cdot}8\ (29)\}$, $B\angle = \{41°{\cdot}7\ (29)\}$. If these measurements are compared with the Fourth and Fifth Dynasties series from Deshasheh and Medum provided by Thomson and MacIver (Table V above) the two will be seen to be remarkably similar. Fourteen characters can be used in calculating the Coefficient of Racial Likeness between them, giving C.L. $= -0{\cdot}50 \pm {\cdot}13$ and C.L. $= -0{\cdot}46 \pm {\cdot}19$ for 6 indices and angles. So, as far as we can tell, the populations were absolutely identical and they only differ from the pure Lower Egyptian type by having a greater basio-bregmatic height.

‡ *Biometrika*, Vol. XVI. pp. 237—239. The series is known as the E series and the ♂ standard deviations had previously been published in *Biometrika*, Vol. XIV. p. 208 and the ♀ in Vol. XVI. p. 14.

from a single cemetery south of Gizeh which was excavated by Sir Flinders Petrie (see (5), p. 29). In Table VIII the measurements of a short series from Sakkarah dating from Ptolemaic times are compared with two other series from Lower Egypt—the 26th to 30th Dynasties from Gizeh and Broca's Fourth Dynasty series from Sakkarah. The individual measurements of the Ptolemaic skulls are in Broca's unpublished note-books and the means are given by Chantre ((15), p. 131) and other French craniologists. They are presented in as unsatisfactory a way as the rest of that material (see footnote ‖ to Table I, p. 15 above).

TABLE VIII.

Mean Measurements of Lower Egyptian Series. Males.*

	L	*B*	*H'*	*B'*	*J*	*DC*
E Series, 26th to 30th Dynasties	185·3 (895)	138·9 (896)	132·4 (884)	94·8 (935)	128·7 (785)	—
4th Dynasty, Sakkarah (Broca) ...	185 (31)	141 (31)	137 (31)	95 (31)	131 (31)	24 (31)
Ptolemaic Period, Sakkarah (Broca)	181 (13)	139 (13)	136 (13)	97 (13)	126 (13)	23 (13)
	O_1'	O_2	*NH'*	*NB*	*fml*	*fmb*
E Series, 26th to 30th Dynasties	38·4 (886)	33·7 (888)	51·5 (898)	24·4 (893)	35·6 (894)	30·0 (905)
4th Dynasty, Sakkarah (Broca) ...	38 (31)	33 (31)	52 (31)	24 (31)	36 (31)	30 (31)
Ptolemaic Period, Sakkarah (Broca)	39 (13)	33 (13)	50 (13)	24 (13)	35 (13)	29 (13)
	100 *B/L*	100 *H'/L*	100 *B/H'*	100 O_2/O_1'	100 *NB/NH'*	100 *fmb/fml*
E Series, 26th to 30th Dynasties	75·1 (884)	71·5 (884)	104·9 (884)	87·8 (886)	47·4 (893)	84·5 (889)
4th Dynasty, Sakkarah (Broca) ...	76·2 (31)	74·1 (31)	102·9 (31)	86·8 (31)	46·1 (31)	83·3 (31)
Ptolemaic Period, Sakkarah (Broca)	76·8 (13)	75·1 (13)	102·2 (13)	84·6 (13)	48·0 (13)	82·9 (13)

* A full account of the measurements of the 26th to 30th Dynasties series is given in the footnote to Table XIV. The indices of Broca's two series in Table VIII were all calculated from the mean values of the component lengths.

The following Coefficients of Racial Likeness were found:

TABLE IX.

Coefficients of Racial Likeness with the E Series (26th to 30th Dynasties) from Gizeh.

	26th to 30th Dynasties (885·4)			
	All Characters		Indices and Angles	
	C. L.	Number of Characters	C. L.	Number of Characters
1st Dynasty Private Tombs (33·6) ...	16·04 ± ·13	14	24·79 ± ·19	6
1st and 2nd Dynasties Royal Tombs (21·8)	1·20 ± ·13	14	1·85 ± ·19	6
1st Dynasty Royal Tombs (Motley) (33·2)	9·75 ± ·09	31	4·25 ± ·14	12
4th Dynasty (Broca) (31·0)	3·89 ± ·13	13	3·62 ± ·24	4
4th and 5th Dynasties (39·9)	3·43 ± ·13	14	3·44 ± ·19	6

The 26th to 30th Dynasties series contains such a large number of skulls that Coefficients between it and closely allied types may be as large as those which would indicate distinct racial differences when shorter series were being compared. All the Coefficients with the Upper Egyptian types (not given in Table IX) are of the same order as those with the First Dynasty Private Tombs series above. The 26th to 30th Dynasties skulls are of the Lower Egyptian type and they show strong affinity with all the representatives of that race except Motley's aberrant series. From the First to the 30th Dynasties the type had hardly changed at all and the latest series shows no progression on that of the Fourth and Fifth Dynasties. A comparison of individual characters confirms that conclusion. The values of the distinctive characters $100\,B/H'$, $100\,NB/NB'$ and $N\angle$ as well as its cephalic index and calvarial breadth all relegate the 26th to 30th Dynasties series to the Lower Egyptian group and there is no suggestion of there having been any admixture whatever between the more stationary population of the Fayum and the Upper Egyptian race. If the E series is intercalated between the First and Fourth Dynasties series we get a more or less orderly sequence for the characters B and $100\,B/L$ but the increase in calvarial breadth is apparently not associated with an increase in zygomatic breadth.

	B	$100\,B/L$	J	H'	$100\,B/H'$	$100\,H'/L$
1st Dynasty Royal Tombs (Motley) ...	136·2	73·8	129·7	132·1	103·6	71·5
1st and 2nd Dynasties Royal Tombs ...	137·7	74·8	129·9	132·8	103·7	72·0
26th to 30th Dynasties...	138·9	75·1	128·7	132·4	104·9	7[illegible]
4th Dynasty (Broca)	137	76·2	131	137	102·9	74·1
4th and 5th Dynasties	139·3	75·4	127·1	136·0	102·5	72·6
Ptolemaic Period (Broca)	139	76·8	126	136	102·2	75·1

The basio-bregmatic height (H') and the indices $100\,B/H'$ and $100\,H'/L$ affected by it are responsible for all the greatest values of α between the six series. There is a clear distinction for those characters between the upper and lower triplets. This is somewhat surprising but it cannot disturb our belief in the very close affinities of all the types. There is clear evidence to show that the pure Upper Egyptian type is represented by the 26th to 30th Dynasties series*, and hence we must suppose that the other has been modified by some process of evolution or racial admixture. As far as we can tell that process only affected the calvarial height so that the only significant differences between the Fourth and Fifth Dynasties series and the 26th to 30th are for H', $100\,B/H'$, $100\,H'/L$ and NB, the last difference being small. We may note that the skulls of the First Dynasty series, which represent the Lower Egyptian type in Upper Egypt, are not absolutely pure, but they incline slightly towards the Upper Egyptian type, as shown in Fig. 1. That arrangement is justified by a comparison of all the Coefficients.

* The transformed Upper Egyptian types incline towards the 26th to 30th Dynasties series for all characters.

6. *Upper Egyptian Series later than the Fifth Dynasty measured by Thomson and MacIver.* There are six of these series and the mean measurements, which have not previously been published, are provided in Table XI. The number of skulls in each is sufficiently large to give a reliable determination of type. From a cursory inspection of the figures it is clear that three of the series resemble one another closely and are very similar to the First Dynasty Private Tombs skulls. The Coefficients of Racial Likeness are given in Table X.

TABLE X.

Coefficients of Racial Likeness between Series of Early and Middle Dynastic Crania.

		1st Dynasty Private Tombs (33·6)	6th to 12th Dynasties (168·6)	12th to 15th Dynasties (65·9)	18th Dynasty Shekh Ali (40·0)
1st Dynasty ... ... Private Tombs (33·6) ...	14 Characters ...	—	2·88 ± ·13	1·88 ± ·13	3·42 ± ·13
	6 Indices and Angles	—	−0·13 ± ·19	0·87 ± ·19	1·58 ± ·19
6th to 12th Dynasties (168·6)	14 Characters ...	2·88 ± ·13	—	2·54 ± ·13	2·87 ± ·13
	6 Indices and Angles	−0·13 ± ·19	—	2·84 ± ·19	3·21 ± ·19
12th to 15th Dynasties (65·9)	14 Characters ...	1·88 ± ·13	2·54 ± ·13	—	0·14 ± ·13
	6 Indices and Angles	0·87 ± ·19	2·84 ± ·19	—	−0·25 ± ·19
18th Dynasty Shekh Ali (40·0)	14 Characters ...	3·42 ± ·13	2·87 ± ·13	0·14 ± ·13	—
	6 Indices and Angles	1·58 ± ·19	3·21 ± ·19	−0·25 ± ·19	—

Paying due regard to the differing numbers of skulls in the series, we may conclude from the Coefficients that they are arranged by their relationships in a sequence which is identically the same as the chronological order. This suggests that the skull form was being gradually modified in a particular direction though the change from the First to the 18th Dynasties was quite slight. Significant differences are only shown for the characters $100\,NB/NH'$, J, H', NB, $G'H$ and $100\,H'/L$. The last three of these are characters which are constant for pure Egyptian types, but no real significance can be attached to their small and erratic differences in the case of the four Early Dynastic series. There is no gradation in calvarial breadths or cephalic indices, but $100\,B/H'$, $100\,N/H'$, H' and $N\angle$ all give orders which are almost precisely the same as the one determined from the Coefficients and the series are obviously links in the transition from the Upper to the pure Lower Egyptian type of the 26th to 30th Dynasties series.

	$100\,B/H'$	H'	$100\,NB/NH'$	$N\angle$*
1st Dynasty Private Tombs ...	99·0	135·1	50·4	66°·2
6th to 12th Dynasties ...	99·6	134·8	50·4	65°·7
12th to 15th Dynasties ...	100·6	133·0	48·8	65°·7
18th Dynasty (Shekh Ali) ...	101·1	132·4	48·0	65°·6
26th to 30th Dynasties ...	104·9	132·4	47·4	63°·9

* Part of the difference between the last series and the others for this character may be due to differences in the determination of the alveolar point. See p. 53 below.

TABLE XI.

Upper Egyptian Series later than the 5th Dynasty measured by Thomson and MacIver. Males.*

Dynasty	6th—12th	12th—15th	18th	18th	Ptolemaic Period	Roman Period
Locality	Denderah	Hou and Abydos	Shekh Ali	Abydos	Denderah	Denderah
F	181·4 (175)	181·3 (67)	180·7 (41)	183·2 (50)	181·9 (78)	179·3 (50)
L	182·5 (175)	183·0 (67)	182·7 (41)	185·8 (50)	183·5 (77)	181·4 (50)
B	134·3 (173)	133·8 (66)	133·8 (41)	137·3 (50)	135·3 (78)	136·4 (50)
H'	134·8 (171)	133·0 (66)	132·4 (41)	134·2 (50)	131·5 (77)	130·7 (50)
LB	100·8 (167)	100·6 (66)	100·4 (40)	101·5 (50)	100·7 (76)	99·9 (50)
$G'H$	69·9 (169)	71·0 (66)	69·0 (40)	71·8 (50)	71·0 (76)	70·5 (50)
J	124·5 (157)	126·4 (61)	125·1 (37)	128·9 (49)	128·6 (74)	127·5 (46)
NH'	51·2 (168)	51·2 (67)	50·8 (39)	52·9 (50)	52·1 (77)	51·6 (48)
NB	25·8 (167)	25·0 (67)	24·4 (39)	25·8 (50)	25·7 (77)	25·5 (48)
GL	96·5 (166)	96·7 (66)	95·8 (40)	95·5 (50)	95·0 (76)	94·2 (50)
$100\ B/L$	73·6 (173)	73·1 (66)	73·2 (41)	73·9 (50)	73·7 (77)	75·2 (50)
$100\ H'/L$	73·9 (171)	72·7 (66)	72·5 (41)	72·2 (50)	71·7 (77)	72·1 (50)
$100\ B/H'$	99·6 (171)	100·6 (66)	101·1 (41)	102·3 (50)	102·9 (77)	104·4 (50)
$100\ (B-H')/L$	−0·3 (171)	0·4 (66)	0·8 (41)	1·7 (50)	2·1 (77)	3·1 (50)
$100\ NB/NH'$	50·4 (167)	48·8 (67)	48·0 (39)	48·8 (50)	49·3 (77)	49·4 (48)
$N\angle$	65°·7 (166)	65°·7 (66)	65°·6 (40)	63°·9 (50)	64°·3 (76)	64°·3 (50)
$A\angle$	72°·8 (166)	72°·0 (66)	73°·2 (40)	73°·3 (50)	73°·3 (76)	73°·2 (50)
$B\angle$	41°·5 (166)	42°·3 (66)	41°·2 (40)	42°·8 (50)	42°·4 (76)	42°·5 (50)

* All the indices and angles in this table were calculated from the mean values of the component lengths.

TABLE XII.

Coefficients of Racial Likeness between series of the Lower Egyptian Type.

		18th Dynasty Abydos (49·9)	Ptolemaic* Period (76·6)	Roman Period (49·3)	1st and 2nd Dynasties Royal Tombs (21·8)	4th and 5th Dynasties (39·9)	26th to 30th Dynasties (885·4)
18th Dynasty Abydos (49·9)	14 Characters ...	—	0·89 ± ·13	2·92 ± ·13	3·16 ± ·13	0·87 ± ·13	5·88 ± ·13
	6 Indices and Angles	—	−0·56 ± ·19	1·15 ± ·19	0·51 ± ·19	0·61 ± ·19	5·40 ± ·19
Ptolemaic Period (76·6)	14 Characters ...	0·89 ± ·13	—	0·72 ± ·13	0·64 ± ·13	4·26 ± ·13	10·05 ± ·13
	6 Indices and Angles	−0·56 ± ·19	—	1·31 ± ·19	0·93 ± ·19	2·00 ± ·19	8·68 ± ·19
Roman Period (49·3)	14 Characters ...	2·92 ± ·13	0·72 ± ·13	—	0·65 ± ·13	3·55 ± ·13	5·35 ± ·13
	6 Indices and Angles	1·15 ± ·19	1·31 ± ·19	—	0·10 ± ·19	0·48 ± ·19	2·19 ± ·19

* Between the short Naqada B, T and R Series (see above, p. 18) and Thomson and MacIver's Ptolemaic we find coefficients of 0·04 ± ·13 for 14 characters and −0·88 ± ·19 for 6 indices and angles. It is possible that the skulls that make up the former series represented a population of almost pure Lower Egyptian type which was intrusive in Upper Egypt in early times and had not mixed with the native population there. That the two Naqada groups were almost certainly contemporaneous is shown by the similarity of their grave furniture.

A comparison of the Coefficients of Racial Likeness with the 26th to 30th Dynasties series confirms the conclusion already arrived at that as the dynasties advanced the population of Upper Egypt approximated more and more to the stable type of the Lower Egyptian people.

The series of which the means are shown in the last three columns of Table XI appear to be of almost pure Lower Egyptian type although they came from Upper Egypt. The Coefficients between them and with three other series to which they are closely related are given in Table XII. The Ptolemaic and 18th Dynasty Abydos series are almost identical, while the Roman is very similar to the former but distinct from the latter. All three are closer to the First and Second Dynasties Royal Tombs series, than to any other of the Lower Egyptian type and the skulls of the 18th Dynasty from Abydos are the only ones that show very close affinity with those of the Fourth and Fifth Dynasties from Deshasheh and Medum in Lower Egypt.

The Fourth and Fifth Dynasties series was shown to be of absolutely pure Lower Egyptian type for all characters except the basio-bregmatic height and the index $100\,B/H'$. The 18th Dynasty Abydos series, though still inclining slightly towards the Upper Egyptian stock, is also primarily of the Lower Egyptian type and it too is characterised by a large H'. We can only suppose that there was a very slight alien element in both those populations. We may intercalate them between the other series to give the more or less orderly sequences shown below for all characters except H' and B/H', but the arrangement in Fig. 1, p. 2, more accurately represents their relationships.

	B	$100\,B/L$	$100\,B/H'$	$100\,NB/NH'$	H'
Ptolemaic Period	135·3	73·7	102·9	49·3	131·5
18th Dynasty Abydos	137·3	73·9	102·3	48·8	134·2
Roman Period	136·4	75·2	104·4	49·4	130·7
1st and 2nd Dynasties Royal Tombs ...	137·7	74·9	103·7	48·5	132·8
26th to 30th Dynasties	138·9	75·1	104·9	47·4	132·4
4th and 5th Dynasties	139·3	75·4	102·5	48·2	136·0

There is a considerable difference between the 18th Dynasty skulls from Abydos and the contemporary series from Shekh Ali, the Coefficients between the two being $4{\cdot}87 \pm {\cdot}13$ for 14 characters and $0{\cdot}73 \pm {\cdot}19$ for six indices and angles. The Fourth and Fifth Dynasties and the 18th Dynasty Abydos series are the only two measured by Thomson and MacIver which do not represent intermediate types between the two races which populated Upper and Lower Egypt in the earliest times of which we have record, without the slightest infusion, as far as we can tell, of racial elements foreign to Egypt.

7. *Upper Egyptian Series later than the Fifth Dynasty measured by Chantre and Broca**. This material has all the defects of Chantre's prehistoric series (see

* We are unable to make any use of two papers by Pruner-Bey (17) and Biasutti (18) that are frequently referred to. The total number of skulls dealt with in each is considerable, but when divided into geographical and time groups they become quite inadequate.

TABLE XIII.

Mean Measurements of Upper Egyptian Series later than the 5th Dynasty measured by Chantre and Broca.*

Period	11th Dynasty	18th Dynasty	18th Dynasty	26th Dynasty	Ptolemaic
Locality	Thebes (Drah abou'l Neggah)	Gournah	Thebes	Gournah	Gournah
Measured by	Broca	Chantre	Broca	Chantre	Chantre
Number of Skulls	18	20	18	20	25
L	184	184	182	182	180
B	136	135	136	136	137
H'	129	131	133	—	—
B'	93	95	91	97	94
J	130	129	126	130	127
DC	29	25	26	23	23
O_1'	37	38	38	40	39
O_2	33	32	31	33	32
NH'	51	52	49	52	49
NB	24	26	25	26	25
fml	35	34	34	34	34
fmb	29	29	29	29	30
U	—	—	510	—	—
$100\,B/L$	73·9	73·4	74·7	74·7	76·1
$100\,H'/L$	70·1	71·2	73·1	—	—
$100\,B/H'$	105·4	103·1	102·3	—	—
$100\,(B-H')/L$	3·8	2·2	1·6	—	—
$100\,O_2/O_1'$	89·2	86·8	81·6	82·5	81·1
$100\,NB/NH'$	47·1	50·0	51·0	50·0	51·0
$100\,fmb/fml$	82·9	85·3	85·3	85·3	88·2

* All the indices in this Table were calculated from the mean values of the component lengths.

footnote ‖ to Table I, p. 14). Ignoring the shorter series, there are five available ranging from the 11th Dynasty to Ptolemaic times. The mean measurements given in Table XIII were taken from various tables in (15) and (16), p. 507. All the series appear to be intermediate between the Upper and Lower Egyptian types: the breadths and cephalic indices arrange them in orders which are almost identically the same as the chronological order. As far as we can tell all the series are pure Egyptian and their relationships seem to confirm the hypotheses that were deduced from more adequate data.

8. *Upper Egyptian Series later than the Fifth Dynasty measured by German Craniometricians.* These are of very considerable importance because the majority are long series and an adequate number of measurements is provided for all of them. The individual measurements of two of the most important have been furnished by Dr Emil Schmidt in the Leipzig portion of the German Anthropological Catalogue (19). The longer (*loc. cit.* pp. 57—117) was reduced by Professor Karl Pearson and Dr Alice Lee and the means were first quoted in Fawcett's paper ((6), p. 426). The skulls came from Thebes and belonged to the

TABLE XIV.

Mean Measurements of Middle and late Dynastic Series of Egyptian Crania. Males.

Date	Middle Dynasties	Early and Middle Dynasties	Modern	18th to 21st Dynasties	18th to 20th Dynasties	18th Dynasty	18th and 19th Dynasties	26th to 30th Dynasties
Locality	El-Kubanieh North	El-Kubanieh South	Abyssinian Tigre Dist.	Thebes	Thebes	Thebes?	Abydos	Gizeh
Measured by	Toldt	Toldt	S. Sergi	Schmidt	Stahr	Schultz and Oetteking	Schmidt	Davin and Pearson
C	1355·0 (27)	1366·1 (67)	—	1387·6 (164)	1451·0 (50)	1393·8 (21)	1413·7 (25)	1438·9 (753)
L	182·2 (37)	182·7 (70)	183·7 (65)	181·9 (169)	183·5 (57)	185·1 (25)	183·1 (33)	185·3 (895)
L'	181·8 (37)	182·5 (66)	—	—	—	—	—	184·6 (872)
B	134·2 (37)	135·9 (70)	136·3 (65)	136·6 (169)	137·3 (56)	137·6 (25)	139·2 (33)	138·9 (896)
B'	90·4 (37)	92·7 (70)	94·6 (65)	93·8 (168)	94·9 (59)	95·1 (25)	96·2 (33)	94·8 (935)
H	—	—	—	134·1 (169)	—	134·3 (24)	132·4 (28)	134·1 (884)
H'	135·9 (35)	132·9 (65)	133·5 (64)	131·9 (168)	133·7 (54)	133·9 (25)	130·9 (33)	[132·4 (884)]
OH	116·1 (36)	115·0 (66)	112·6 (65)	114·3 (169)	114·1 (56)	113·8 (25)	113·5 (33)	113·4 (887)
LB	101·6 (34)	100·1 (63)	99·8 (64)	100·6 (169)	101·9 (56)	101·8 (25)	100·7 (33)	101·6 (896)
Q	—	—	309·8 (65)	306·1 (169)	—	307·8 (25)	308·1 (32)	307·2 (878)
S	372·7 (33)	371·8 (69)	372·7 (58)	372·4 (168)	372·1 (51)	374·8 (25)	374·5 (32)	371·9 (884)
S_1	126·0 (36)	128·2 (69)	130·2 (65)	—	125·7 (54)	128·5 (25)	—	127·4 (928)
S_2	130·9 (33)	127·8 (67)	128·9 (65)	—	129·8 (50)	129·6 (25)	—	128·5 (908)
S_3	115·1 (30)	116·0 (66)	113·7 (57)	—	114·7 (54)	117·8 (25)	—	115·2 (889)
S_3'	97·3 (30)	97·8 (66)	95·8 (57)	96·5 (166)	96·1 (54)	97·8 (25)	96·3 (31)	94·7 (885)
U	505·6 (35)	506·2 (69)	513·0 (65)	510·8 (168)	512·6 (54)	512·8 (25)	517·1 (33)	518·7 (876)
G'H	70·1 (29)	70·2 (62)	69·1 (65)	68·8 (169)	70·1 (56)	69·7 (25)	71·3 (30)	70·4 (845)
GB	94·9 (29)	95·2 (57)	93·1 (57)	95·7 (161)	95·5 (54)	95·3 (24)	95·9 (25)	95·2 (877)
J	125·5 (26)	125·7 (54)	125·0 (42)	128·3 (166)	128·6 (53)	130·7 (25)	131·3 (33)	128·7 (785)
NH	—	—	—	50·61 (169)	—	51·1 (25)	52·1 (31)	51·65 (898)
NH'	49·8 (37)	49·4 (65)	48·9 (64)	[50·4 (169)]	50·8 (56)	[50·9 (25)]	[51·9 (31)]	[51·5 (898)]
NB	24·8 (36)	25·0 (63)	25·0 (63)	25·26 (166)	25·6 (53)	25·1 (25)	24·9 (26)	24·36 (893)
DC	23·4 (34)	23·3 (59)	22·0 (65)	—	21·1 (52)	24·3 (11)	24·2 (31)	—
SC	—	—	9·53 (64)	—	10·8 (52)	—	—	—
O_1	40·8 (36)	41·3 (66)	40·0 (65)	—	—	—	—	41·13 (886)
O_1'	38·1 (35)	38·4 (64)	[37·3 (65)]	38·14 (169)	38·3 (59)	38·3 (18)	38·7 (30)	[38·4 (886)]
O_2	32·3 (35)	32·8 (66)	33·2 (65)	32·96 (169)	32·8 (53)	33·0 (24)	33·5 (31)	33·73 (888)
G_1'	46·4 (25)	47·0 (55)	—	46·25 (160)	46·2 (43)	47·5 (23)	46·2 (26)	46·28 (821)
G_2	—	—	—	40·83 (151)	—	41·6 (21)	41·1 (21)	38·81 (703)
GL	96·5 (25)	98·1 (55)	95·2 (64)	96·41 (167)	96·1 (53)	96·0 (25)	96·2 (31)	95·03 (832)
fml	34·9 (32)	34·7 (64)	36·1 (57)	35·10 (168)	35·2 (55)	36·4 (25)	35·9 (31)	35·56 (894)
fmb	29·2 (32)	28·9 (64)	29·5 (59)	29·80 (165)	29·7 (54)	30·7 (25)	30·2 (30)	29·96 (905)
100 *B*/*L'*	73·7 (37)	74·5 (66)	—	—	—	—	—	75·34 (866)
100 *H*/*L*	—	—	—	{73·7 (169)}	—	{72·6 (24)}	73·7 (33)	72·46 (871)
100 *B*/*L*	73·7 (37)	74·4 (70)	{74·2 (65)}	75·07 (169)	{74·8 (56)}	74·4 (25)	76·0 (33)	75·06 (884)
100 *B*/*H*	—	—	—	{101·9 (169)}	—	{102·5 (24)}	103·1 (33)	103·70 (873)
100 *H'*/*L*	74·7 (35)	72·9 (65)	{72·7 (64)}	{72·5 (168)}	{72·9 (54)}	72·3 (25)	71·5 (33)	{71·5 (884)}
100 *B*/*H'*	{98·7 (35)}	{102·3 (65)}	{102·1 (64)}	{103·6 (168)}	{102·7 (54)}	102·8 (25)	106·4 (33)	{104·9 (884)}
100 (*B*−*H'*)/*L*	{−0·9 (35)}	{1·6 (65)}	{1·5 (64)}	{2·6 (168)}	{2·0 (54)}	{1·8 (24)}	{4·5 (33)}	{3·5 (884)}
100 *G'H*/*GB*	{73·9 (29)}	{73·7 (57)}	{74·2 (57)}	72·32 (155)	{73·4 (54)}	{73·1 (24)}	{74·4 (25)}	74·13 (823)
100 *NB*/*NH*	—	—	—	50·07 (166)	—	{49·1 (25)}	47·9 (30)	47·31 (881)
100 *NB*/*NH'*	49·9 (36)	50·8 (63)	{51·1 (63)}	{50·1 (166)}	{50·4 (53)}	{49·3 (25)}	{48·0 (26)}	{47·4 (893)}
100 O_2/O_1	{79·2 (35)}	{79·4 (66)}	{83·0 (65)}	—	—	—	—	82·12 (878)
100 O_2/O_1'	84·9 (35)	85·7 (64)	{89·0 (65)}	86·5 (169)	{85·6 (53)}	{86·2 (18)}	86·8 (30)	{87·8 (886)}
100 *fmb*/*fml*	84·0 (31)	83·4 (63)	{81·7 (57)}	85·50 (165)	{84·4 (54)}	84·4 (25)	84·3 (28)	84·47 (889)
100 G_2/G_1'	—	—	—	88·5 (151)	—	87·6 (19)	89·3 (20)	{83·9 (703)}
P∠	85°·6 (28)	84°·9 (58)	—	85°·5 (47)	85°·2 (55)	85°·9 (23)	—	85°·8 (795)
N∠	{65°·2 (25)}	{67°·5 (55)}	{65°·4 (64)}	{66°·1 (167)}	{64°·5 (53)}	{65°·0 (25)}	{65°·2 (30)}	63°·9 (818)
A∠	{73°·4 (25)}	{71°·0 (55)}	{73°·2 (64)}	{73°·0 (167)}	{74°·0 (53)}	{74°·0 (25)}	{72°·2 (30)}	74°·1 (818)
B∠	{41°·4 (25)}	{41°·5 (55)}	{41°·4 (64)}	{40°·9 (167)}	{41°·5 (53)}	{41°·0 (25)}	{42°·5 (30)}	41°·9 (818)
θ_1	{29°·2 (25)}	{27°·6 (55)}	—	{28°·4 (47)}	{30°·3 (53)}	{28°·8 (23)}	—	30°·03 (795)
θ_2	{12°·2 (25)}	{13°·9 (55)}	—	{12°·5 (47)}	{11°·2 (53)}	{11°·9 (23)}	—	11°·89 (795)
Oc. I.	{62·3 (30)}	{62·0 (66)}	{61·9 (57)}	—	{61·4 (54)}	{60·6 (25)}	—	61·54 (884)
W (gms.)	—	—	605·2 (65)	—	598·4 (48)	—	—	—
EO W	—	—	—	—	—	97·4 (25)	—	—

Notes to Table XIV.

The indices and angles in curled brackets were calculated from the mean values of the component lengths. The lengths H', O_1' and NH' in squared brackets were found from the similar measurements H, O_1R and NH, R on the assumption that the differences would be the same as for Motley's very similar First Dynasty series (in Table V, p. 24, above). Those differences are very similar to values found for other ancient Egyptian series and, as far as we can tell, they are quite constant for all modern races of man. The capacities given by Stahr were probably determined by a method which would make them not comparable with the others. Tildesley (*Biometrika*, Vol. XIII. p. 217) has shown that in taking auricular heights on the Ranke's craniophor with the flute-topped plugs which are at present used in the Biometric Laboratory*, a correction should be made by subtracting 2·5 mm. from the scale readings, and that is confirmed by comparison of mean direct and type contour measurements. That correction has been made to Motley's First Dynasty means in Table V. In the case of the series measured by Sergi and Toldt it is doubtful whether the nasal height measurement corresponds to our NH or NH': the difference between the two is small. Schmidt took the length to the lowest point of the pyriform aperture, i.e. whichever is greater of NH, R and NH, L, so the mean would be very slightly greater than that of either NH, R or NH, L. Oetteking does not say on which side he took the measurement. The Occipital Index is a complex function of the arc S_3 and the chord S_3', so it is not surprising that its mean value calculated from the indices of individual skulls should not be the same as the index found from the mean values of S_3 and S_3'. For all races the former exceeds the latter and the differences are 0·5 for the 26th to 30th Dynasties series, 1·0 for the Naqada A and Q series, 0·9 for the Naqada B, T and R series and 1·1 for Motley's First Dynasty series. It can be shown that the difference is not dependent on the number of skulls in the series so we may accept the mean of these four as a corrective factor for Egyptian types. The Occipital Indices in curled brackets in Table XIV were obtained by adding 0·9 to the indices deduced from the mean values of S_3 and S_3'.

In finding the means of the Abyssinian series, four juvenile skulls (Nos. 7, 13, 53 and 59) were excluded and one juvenile (No. 15) was excluded from Toldt's El-Kubanieh South series. The GL length of Schmidt's 18th to 21st Dynasties series was given as 95·79 (169); omitting one juvenile skull (No. 630) I find that it is 96·41 (167) and the alteration affects the angles of the fundamental triangle.

Several mean measurements of the 26th to 30th Dynasties series other than those in Table XIV were determined: viz. $F=182·8$ (895), $Q=307·5$ (876), $O_1L=40·75$ (880), $O_2L=33·83$ (888) (the O_1, O_2 and $100\,O_2/O_1$ in the table refer to the right orbit), $G_1=49·70$ (817), $100\,H/L'=72·72$ (863), $100\,O_2/O_1$, $L=83·06$ (876), $100\,G_2/G_1=78·36$ (668).

18th to 21st Dynasties. The means of the other series ((19), pp. 64—74), which belonged to the 18th and 19th Dynasties and came from Abydos, are now given for the first time. Schmidt gives quite complete measurements, following in general the *Frankfurter Verständigung*, but a few of them might have been defined with greater precision. The means are given in Table XIV. Both series are predominantly of the Lower Egyptian type and they show very close affinities with some previously considered.

When comparing the Coefficients of Racial Likeness calculated for several series of different lengths it is of the utmost importance to take into consideration the number of skulls on which each series is based. If that is done the five series of Lower Egyptian type shown in the upper part of Table XV are arranged by their inter-relations in the order shown. Schmidt's Abydos type stands very close to those of the 26th to 30th Dynasties and of the Royal Tombs of the First and Second Dynasties, but his longer series from Thebes of the same date is far closer to the Roman Period skulls. The former is of quite pure Lower Egyptian type and the latter diverges slightly from that race in the direction of the Upper Egyptian. The latter point is brought out clearly by a comparison with the five types shown in the lower half of Table XV, from the 18th Dynasty Shekh Ali to the Early

[* Earlier measures were taken with the *pointed* plugs for which the craniophor is scaled, and I believe there is not adequate ground for modifying their values as Tildesley proposed.—ED.]

TABLE XV.

Coefficients of Racial Likeness with Schmidt's Series.

	18th and 19th Dynasties Abydos, Schmidt (29·9)				18th to 21st Dynasties Thebes, Schmidt (162·4)			
	All Characters		Indices and Angles		All Characters		Indices and Angles	
	C. L.	Number of Characters	C. L.	Number of Characters	C. L.	Number of Characters	C. L.	Number of Characters
26th to 30th Dynasties (885·4)	2·01 ± ·09	27	2·80 ± ·16	9	20·55 ± ·09	28	24·57 ± ·15	10
18th and 19th Dynasties Schmidt (29·9)	—	—	—	—	2·49 ± ·09	29	2·41 ± ·15	10
1st and 2nd Dynasties Royal Tombs* (21·8)	0·81 ± ·13	14	1·32 ± ·19	6	1·37 ± ·13	14	0·27 ± ·19	6
Roman Periods* (49·3) ...	1·74 ± ·13	14	1·05 ± ·19	6	1·23 ± ·13	14	0·11 ± ·19	6
18th to 21st Dynasties Schmidt (162·4)	2·49 ± ·09	29	2·41 ± ·15	10	—	—	—	—
1st Dynasty Royal Tombs Motley (33·2)	2·77 ± ·09	29	2·21 ± ·15	10	7·48 ± ·09	30	4·50 ± ·14	11
4th and 5th Dynasties* (39·9)	2·89 ± ·13	14	2·53 ± ·19	6	4·70 ± ·13	14	1·77 ± ·19	6
18th Dynasty Shekh Ali* (40·0)	7·43 ± ·13	14	7·45 ± ·19	6	4·28 ± ·13	14	5·38 ± ·19	6
6th to 12th Dynasties* (168·6)	15·73 ± ·13	14	19·03 ± ·19	6	15·84 ± ·13	14	19·16 ± ·19	6
1st Dynasty Private Tombs (33·6)*	8·90 ± ·13	14	14·98 ± ·13	6	6·77 ± ·13	14	9·80 ± ·19	6
Late Pre-dynastic* (88·5)	13·66 ± ·13	14	19·41 ± ·13	6	14·07 ± ·13	14	22·05 ± ·19	6
Early Pre-dynastic* (35·8)	16·91 ± ·13	14	21·81 ± ·19	6	13·67 ± ·13	14	19·55 ± ·19	6

* These series were measured by Thomson and MacIver.

Pre-dynastic. They represent the earlier transitional stages from the pure Upper Egyptian type to the distinct one of Lower Egypt. The Coefficients with Schmidt's series are all of a higher order than the ones in the upper part of the table, but the order of arrangement is in both cases the same as that which was arrived at in considering the inter-relations of the earlier series themselves. It is quite reasonable to place the 6th to 12th Dynasties series before that of the 1st Dynasty Private Tombs since the disparity between the numbers of skulls in each is so great. Remembering that Schmidt's 18th to 21st Dynasties series is far longer than his other, we may conclude that its relationships with the Upper Egyptian types, though remote, are distinctly closer than are those between the latter and Schmidt's Middle Dynastic series from Abydos. A detailed comparison of individual measurements need hardly be made; it would confirm the conclusions that have already been reached. The greatest values of α are almost invariably found between the characters $100\,B/H'$, B, $100\,B/L$ and L, while H', $N\angle$ and $100\,NB/NH'$ show differences that are smaller but still significant in some cases; and certain other characters are reasonably constant throughout. Nevertheless Schmidt's L and $G'H$ of the 18th to 21st Dynasties series are as small as for any other dynastic Egyptian

series and their values differ significantly from those of some closely related types. This alone would be very meagre evidence of alien admixture, but it will be shown later that there are better reasons for supposing that the series in question had diverged slightly from the pure Egyptian stock.

In 1919 C. Toldt (20) furnished individual measurements of Egyptian skulls that had been disinterred at El-Kubanieh near Assuan. The site is 100 miles from Thebes and is more southerly than any other with which we are at present concerned. The whole material came from two graveyards and it was divided into the following groups (see (20), pp. 526, 635 and 636):

El-Kubanieh South { Early Dynastic 62 skulls. / Middle Dynastic 52 skulls.

El-Kubanieh North { Middle Dynastic 31 skulls. / Nomads 26 skulls.

The nomads in the second group were supposed to be related to tribes of the northern Sudan and they apparently belonged to Middle Dynastic times. Unfortunately it is not possible to separate them from the other skulls from the El-Kubanieh North cemetery. The means of the two groups reduced from Toldt's individual measurements are given in Table XIV. The Coefficients between them and the most closely related series are in Table XVI.

TABLE XVI.

Coefficients of Racial Likeness with Toldt's Series from El-Kubanieh.*

		6th to 12th Dynasties (168·6)	El-Kubanieh North (33·1)	12th to 15th Dynasties (65·9)	El-Kubanieh South (63·5)	18th Dynasty Shekh Ali (40·0)
El-Kubanieh South (63·5)	14 Characters ...	5·90 ± ·13	3·24 ± ·13	2·95 ± ·13	—	2·86 ± ·13
	6 Indices and Angles	7·78 ± ·19	6·33 ± ·19	4·61 ± ·19	—	5·30 ± ·19
El-Kubanieh North (33·1)	14 Characters ...	0·81 ± ·13	—	1·41 ± ·13	3·24 ± ·13	1·69 ± ·13
	6 Indices and Angles	−0·14 ± ·19	—	2·60 ± ·19	6·33 ± ·19	2·67 ± ·19

* All the Coefficients in Table XVI are calculated for the 14 characters, of which 6 are indices and angles, available for all Thomson and MacIver's series. Between the two El-Kubanieh series 28 characters can be used giving a Coefficient of 1·60 ± ·09 and of 3·16 ± ·14 for the 11 indices and angles.

It is clear that the El-Kubanieh North and South series are very similar to the Middle Dynastic people of Upper Egypt: they take up the positions shown in Table XVI and Fig. 1. The former is the later in time and yet it has been less changed from the Upper Egyptian type of the First Dynasty than the other and the difference is doubtless to be attributed to the nomadic element. In calculating the Coefficients in Table XVI the greatest values of α were found for the characters $100\,B/H'$, B, $100\,B/L$, H', $N\angle$, $100\,H'/L$ and NH'. The prominence of the last two is rather exceptional. The El-Kubanieh North series, but not the other, has

an unusually large basio-bregmatic height and height index, but other characters are not distinctive from other Middle Dynastic types. The Coefficients between Toldt's two series and Schmidt's Theban skulls of the 18th to 21st Dynasties are given in Table XVII.

TABLE XVII.

Coefficients of Racial Likeness between Schmidt's 18th to 21st Dynasties and Toldt's El-Kubanieh Series.

	18th to 21st Dynasties Schmidt (162·4)			
	All Characters		Indices and Angles	
	C. L.	Number of Characters	C. L.	Number of Characters
El-Kubanieh North (33·1) ...	4·50 ± ·09	27	6·14 ± ·15	10
El-Kubanieh South (63·5) ...	2·65 ± ·09*	27	3·27 ± ·15*	10
18th Dynasty Shekh Ali (40·0) ...	4·28 ± ·13	14	5·38 ± ·19	6

* When these Coefficients are computed for the 14 characters, of which 6 are indices and angles, available for Thomson and MacIver's series, they are increased to 3·51 ± ·13 and 4·23 ± ·09 respectively.

That the El-Kubanieh South series should be closer to Schmidt's 18th to 21st Dynasties than the El-Kubanieh North is quite consistent with the relationships already found, but that it should also be closer than the 18th Dynasty series from Shekh Ali is unexpected. This discordance might be attributed to the fact that Toldt's series was heterogeneous in respect to time, but it will be shown later that a more probable explanation is that both it and Schmidt's 18th to 21st Dynasties series were intermixed very slightly with an alien stock. The positions shown in Fig. 1, p. 2, of the El-Kubanieh relative to the other series are the most reasonable.

We regret that it has only been possible to make slight use of Dr Bruno Oetteking's paper (22) on ancient Egyptian crania. He has not sexed his material and therefore it cannot be used for comparative purposes. It is strange to find a craniological study as late as 1909 in which the writer is content to take means and variations on a population without distinction of sex. The Egyptian skulls measured by Dr Oetteking are in the University of Zürich and in 1918 Dr Adolf Schultz published measurements of 35 of them (25 ♂ and 10 ♀) for the purpose of a special enquiry ((23), measurements on pp. 72—79). Accepting that sexing, the means of the 25 specimens were reduced from Schultz and Oetteking's tables and they are given in Table XIV. The material was probably Theban of the 18th Dynasty. The Coefficients of Racial Likeness are given in Table XVIII.

The short Zürich series is of almost identically the same type as Thomson and MacIver's of the First and Second Dynasties. It takes up a position between the latter and Schmidt's series of the 18th and 19th Dynasties (see Fig. 1, p. 2) and it gives another representation of the wide-spread population of Lower Egyptian type which was predominating in Upper Egypt in Late Dynastic times. There is not the slightest suggestion of alien admixture.

TABLE XVIII.

Coefficients of Racial Likeness with Schultz and Oetteking's Zürich Series.

	Zürich Series (23·9)			
	All Characters		Indices and Angles	
	C. L.	Number of Characters	C. L.	Number of Characters
18th and 19th Dynasties Schmidt (29·9)	0·51 ± ·09	29	1·25 ± ·15	10
1st and 2nd Dynasties Royal Tombs (21·8)	0·13 ± ·13	14	0·25 ± ·19	6
26th to 30th Dynasties (885·4)	2·02 ± ·09	31	1·84 ± ·14	12

In 1907 Dr Hermann Stahr (21) published full individual measurements of 113 skulls from Thebes belonging to the period from the 18th to 20th Dynasties. The means of the 56 skulls given in Table XIV were reduced from Stahr's tables; not a single one is given in the original paper. The series appears to be quite a homogeneous one and it is of almost pure Lower Egyptian type. The Coefficients of Racial Likeness between it and the series which it most closely resembles are given in Table XIX.

TABLE XIX.

Coefficients of Racial Likeness with Stahr's 18th to 20th Dynasties Series.

	18th to 20th Dynasties Stahr (54·1)			
	All Characters		Indices and Angles	
	C. L.	Number of Characters	C. L.	Number of Characters
Roman Period (49·3)	1·42 ± ·13	14	0·62 ± ·19	6
1st and 2nd Dynasties Royal Tombs (21·8)	0·91 ± ·13	14	1·44 ± ·19	6
18th Dynasty Schultz and Oetteking (23·9)	− 0·06 ± ·09	28	− 0·54 ± ·14	11

The bond between Stahr's series and that measured by Schultz and Oetteking is closer than that between any other two series dealt with in this paper. That is not surprising as they were contemporaneous and both came from Thebes.

For the series dealt with in the Biometric Laboratory and those measured by German craniometricians many more characters are available than the 21 provided by Thomson and MacIver. The significances of the differences of the additional characters that are used in the computation of the Coefficients of Racial Likeness can be gauged in the usual way by considering the values of α. No very detailed comparison need be made here: it entirely confirms the conclusions already arrived at. In the following list (Table XX) the types are arranged in order from the almost pure Upper Egyptian type represented by the Naqada crania to the pure Lower Egyptian type of the 26th to 30th Dynasties series. The intermediate series, all of which

came from Upper Egypt, represent the more or less gradual transition from one extreme to the other. Where a character does not differ significantly between the Upper and Lower Egyptian types it is always found that no one of the intervening types differs significantly from either of the extremes. Where, however, the characters of the Upper Egyptian type do differ significantly from those of the Lower then, for those characters, there is a more or less gradual transition from one extreme to the other when the series are arranged in the order shown, which is the order in which the Coefficients of Racial Likeness arrange them. For some characters Motley's First Dynasty series is quite clearly differentiated from the others.

The length characters S, S_1, S_2, S_3, S_3', G_1' and *fml*, the facial breadths GB and the foraminal indices, profile angles and occipital indices of the Upper and Lower Egyptian and all intermediate types seem to be identically the same. Small differences are found between the characters associated with facial heights (O_2, $100O_2/O_1'$ and $100G'H/GH$) and greater differences between those influenced by change in calvarial breadth (B', U, Q' and C), and the gradual transition from the Upper to the Lower Egyptian type can be traced more or less clearly in the case of the latter. In general a comparison of these additional characters entirely confirms the relationships of the two extreme types already noted. The Lower Egyptian race is characterised by its much greater calvarial and zygomatic breadths and cephalic index and its slightly greater facial heights, while all lengths in the sagittal plane and facial breadths other than the zygomatic breadth are the same for the two races. Some mean measurements of the First Dynasty skulls of the Royal Tombs from Abydos are given in Table XX. It has been shown that that type was primarily Lower Egyptian, though inclining slightly towards the Upper

TABLE XX.

Comparison of Additional Mean Measurements.

Series	Constant Characters				Characters showing significant differences					
	S	G_1'	GB	$O_1'R$	B'	U	O_2R	$100\frac{G'H}{GB}$	$100\frac{O_2}{O_1'}$	C
Naqada *A* and *Q* Series ...	372·6	48·0	95·9	39·1	92·6	510·4	32·2	72·7	82·4	1388·6
El-Kubanieh North	372·7	46·4	94·9	38·1	90·4	505·6	32·3	73·9	84·9	1355·0
El-Kubanieh South	371·8	47·0	95·2	38·4	92·7	506·2	32·8	73·7	85·7	1366·1
18th to 21st Dynasties Schmidt	372·4	46·3	95·7	38·1	93·8	510·8	33·0	72·3	86·5	1387·6
18th to 20th Dynasties Stahr ...	372·1	46·2	95·5	38·3	94·9	512·6	32·8	73·4	85·6	1451·0
18th Dynasty Schultz and Oetteking	374·8	47·5	95·3	38·3	95·1	512·8	33·0	73·1	86·2	1393·8
18th and 19th Dynasties Schmidt	374·5	46·2	95·9	38·7	96·1	517·1	33·5	74·4	86·8	1413·7
26th to 30th Dynasties ...	371·9	46·3	95·2	38·4	94·8	518·7	33·7	74·1	87·8	1438·9
1st Dynasty Royal Tombs Motley	369·0	49·5	97·3	40·0	93·0	513·5	33·5	76·8	86·1	1364·3
Modern Abyssinian, Tigre district	372·7	—	93·1	—	94·6	513·0	33·2	74·2	—	—

Egyptian type. The majority of its measurements, including all those of the calvaria, are intermediate between the measurements of the Upper and Lower Egyptian races, but others clearly differentiate the First Dynasty skulls from all other Egyptian types. The greater facial heights $G'H$ and NH', but not O_2, are the most distinctive characters, but $B\angle$, GB, $O_1'R$, G_1' and G_2 are also greater than for any other series and the indices $100\,G'H/GB$ and $100\,fmb/fml$ diverge from the normal type. Motley's series is undoubtedly not pure Egyptian, but we are at present unable to say what other race took part in its make up; it was certainly not negroid.

In the last row of Table XX there are some mean measurements of modern Abyssinian skulls; the complete measurements are in Column 4 of Table XIV. The data are provided by Dr Sergio Sergi (26). The series with which we are concerned came from the extreme north of the country—the Tigre district—and it is distinctly different from two other shorter ones from Central Abyssinia and Somali respectively*. There is a close relationship between the Tigre skulls and Middle Dynastic Egyptian types.

When it is remembered that an interval of well over 3000 years must have separated the Abyssinians from any of the dynastic skulls with which they are compared in Table XXI, the Coefficients appear to be surprisingly low. The bond

TABLE XXI.

Coefficients of Racial Likeness with Modern Abyssinians (Tigre District).

		18th to 21st Dynasties Schmidt (162·4)	Ptolemaic Period (76·6)	18th Dynasty Shekh Ali (40·0)	El-Kubanieh South (63·5)	12th to 15th Dynasties (65·9)
Modern Abyssinian (Tigre District) (61·8)	14 Characters ...	3·48 ± ·13	6·02 ± ·13	2·10 ± ·13	1·16 ± ·13	3·72 ± ·13
	6 Indices and Angles	1·76 ± ·19	2·01 ± ·19	2·49 ± ·19	3·27 ± ·19	3·24 ± ·19
	23 Characters ...	3·57 ± ·10	—	—	1·87 ± ·10	—
	8 Indices and Angles	4·12 ± ·17	—	—	2·46 ± ·17	—

with the El-Kubanieh South series is at least as close as that between the Moorfields and Whitechapel English series of 17th century Londoners†. At the same time we should have anticipated even lower Coefficients if the Abyssinians

* There is not the slightest suggestion of negroid admixture in the series from Central Abyssinia and Somali.

† These series are in *Biometrika*, Vol. v. p. 92. The Coefficients between the Whitechapel (87·2) and Moorfields (24·2) English are 2·05 ± ·09 for 26 characters and 1·73 ± ·16 for 9 indices and angles. We have good reason to suppose that the Coefficients between the Abyssinian and El-Kubanieh South series should be even lower than the values shown. The greatest differences between them are for $N\angle$ and $A\angle$ and for these characters the modern series is similar to all middle dynastic ones and it is very probable that the aberrant values given by Toldt are due to a difference in method of measurement. If $N\angle$ and $A\angle$ are excluded the two series have Coefficients of 0·95 ± ·10 for 21 characters and − 0·13 ± ·19 for 6 indices and angles.

were of pure ancient Egyptian type. In Table XXI the series are arranged in the order given by their own inter-relation but their degrees of resemblance to the Abyssinians do not arrange them in the same order. The Coefficients with Schmidt's 18th to 21st Dynasties series are decidedly lower than we should have anticipated. It is probable that both that series and the El-Kubanieh South were slightly mixed with an alien stock which had a greater effect on the later Abyssinians. The only individual characters which distinguish the latter from the normal middle dynastic type of Upper Egypt are its small *NH'*, *G'H*, *LB* and *GB*, and the nasal index being peculiarly large and the foraminal index peculiarly small. And for those characters the El-Kubanieh South series and Schmidt's of the 18th to 21st Dynasties show a tendency to diverge from the normal in the direction of the Abyssinian values: the *NH'* and 100 *fmb*/*fml* of the former and the *G'H* of the latter are smaller than for any other dynastic series*.

9. *Type Contours of First Dynasty Skulls from Royal Tombs at Abydos.* The horizontal, transverse and sagittal type contours of the First Dynasty skulls from royal tombs at Abydos (see p. 22 above) were constructed by Mr G. H. Motley from the mean measurements of contours of the individual skulls in the manner commonly practised in the Biometric Laboratory†. The mean measurements are given in Tables XXII, XXIII and XXIV and the outlines in Figs. 3—8. The ♂ contours

TABLE XXII.

Mean Measurements of the Transverse Vertical Contours of 32 ♂ *and* 10 ♀ *First Dynasty Skulls from Royal Tombs at Abydos.*

Sex	*MA*	1*R*=1*L*	*M*¼*R*	*M*¼*L*	2*R*	2*L*	3*R*	3*L*	4*R*	4*L*	5*R*	5*L*	6*R*	6*L*
♂	110·6	57·1	60·0	59·3	60·5	60·1	62·9	62·7	65·3	64·6	66·1	64·9	66·0	64·7
♀	108·2	54·3	57·1	56·5	58·1	57·7	60·9	60·4*	63·1	62·4*	64·3	62·9	64·4	62·9

Sex	7*R*	7*L*	8*R*	8*L*	9*R*	9*L*	10*R*	10*L*	*A*¼*R*	*A*¼*L*	*ZR*.*R*		*ZR*.*L*	
											y	*x*	*y*	*x*
♂	65·1	63·3	61·6	59·2	54·0	50·5	39·4	36·0	19·2	17·8	60·2	3·1	59·8	3·7
♀	63·0	60·9	59·7	56·4	52·0	48·0	37·5	34·4	18·9	17·6	57·4	2·6	57·3	3·1

* These lengths are the means of 9 contours.

may be compared with the types of 100 ♂ skulls from Lower Egypt which formed part of the long 26th to 30th Dynasties series (see p. 29 above) and of which the three type contours were constructed by Benington in 1911 (*Biometrika*, Vol. VIII.

* The Abyssinian and the El-Kubanieh South series are the only two with which we are dealing to have a greater frontal (S_1) than parietal arc (S_2): see Table XIV. These measurements are unfortunately not available for Schmidt's series.

† For a full account of the method used see *Biometrika*, Vol. XIV. pp. 227—241.

pp. 173—177). The sagittal contour of the latter series was only given for the arc above the nasio-gamma base line, so there is no representation of the facial bones and base of the skull. On superposing the outlines, the differences between them are seen to be very small: the type zones would actually overlap for the whole of their lengths, so we cannot say that any of the differences are significant. This conclusion is in accordance with that derived from a comparison of direct measurements as the only significant differences which could be distinguished in that way were those between facial measurements.

TABLE XXIII.

Mean Values of the Horizontal Contours of 32 ♂ *and* 10 ♀ *First Dynasty Skulls from Royal Tombs at Abydos.*

Sex	*FO*	*F* ¼*R*	*F* ¼*L*	*F* ½*R*	*F* ½*L*	2*R*	2*L*	2 ½*R*	2 ½*L*	3*R*	3*L*	4*R*	4*L*	5*R*	5*L*
♂	183·9	23·0	23·5*	34·7	35·5†	45·9	47·2‡	46·1	46·8*	47·9	48·0†	53·7	53·8§	61·0	60·8*
♀	177·5	20·7	21·9	32·9	34·6	44·3	45·8	44·1	45·0	46·0	46·6	51·9‡	51·9	58·9	58·8

Sex	6*R*	6*L*	7*R*	7*L*	8*R*	8*L*	9*R*	9*L*	10*R*	10*L*	*O* ¼*R*	*O* ¼*L*	*T*(*R*)*x*	*T*(*L*)*x*	*T*(*R*)*y*	*T*(*L*)*y*
♂	65·8	65·2*	66·9	66·6*	64·3	64·0	57·7	57·8	45·1	45·8	24·4	24·9	20·5	19·9†	47·9	48·9†
♀	63·6	63·9	65·6‡	65·8	63·4‡	64·3	56·9	58·4	45·9	46·4	25·5	25·0	19·9	19·7	45·9	47·2

* These lengths are the means of 31 contours.
† ,, ,, ,, 30 ,,
‡ ,, ,, ,, 9 ,,
§ ,, ,, ,, 29 ,,

TABLE XXIV.

Mean Measurements of the Sagittal Contours of 31 ♂ *and* 10 ♀ *First Dynasty Skulls from Royal Tombs at Abydos.*

Sex	*N*γ	Ordinates above *N*γ													
		O=*N*	*N*¼	1	2	3	4	5	6	7	8	9	γ¾	γ½	γ=10
♂	182·8	15·4	31·9	56·3	69·3	76·7	80·8	82·4	83·0	80·0	70·4	51·4	25·8	18·8	5·27
♀	176·5	15·2	36·2	56·5	68·4	74·6	78·3	80·2	80·0	76·3	67·3	46·2	23·3	16·2	−1·52*

* The negative sign here indicates that the 10th ordinate meets the outline below the *N*γ line on the type.

TABLE XXIV—(*continued*).

Sex	Ordinates below $N\gamma$								Vertex		β		Glabella	
	$O=N$	$N\frac{1}{4}$	1	2	8	9	$\gamma\frac{1}{2}$	$\gamma\frac{1}{4}$	x from N	y	x from N	y	x from N	y
♂	70·7*	67·2	57·4	54·5†	46·1	35·1	26·3	19·0	103·8	83·4	75·0	81·1	1·2	9·7
♀	65·9	61·5	52·6	49·7	48·3	38·4	30·4	23·3	95·7	81·1	75·3	79·0	1·0	9·1

Sex	λ		Sub-Orb.		Aur.		Op.		Inion		Basion		Alv.	
	x from γ	y	x from N	y	x from γ	y	x from γ	y	x from γ	y	from γ	from N	from N	from Bas.
♂	5·7	27·7	10·9	29·1	94·1	28·1	57·0*	53·3*	14·3	32·1	106·6	101·5	74·4	95·7
♀	4·8	22·2	9·8	28·7	90·9	27·9	56·0	55·1	14·3	35·5	102·5‡	98·7‡	70·1‡	93·4‡

Sex	Nose			Palate				Sp.		N. S.		Max. Sub. of Alv. N. S. Chord	
				P'		P							
	(i)	(ii)	(iii)	x from N	y	x from Alv.	y	x from N	y	x from N	y	x from N	y
♂	1·02†	3·0†	6·3*	49·8†	51·5†	36·1†	19·6†	67·0	34·0	6·4	54·1	3·4	62·5
♀	0·82	2·9	6·7	48·0	48·0	38·2	18·6	64·6‡	32·8‡	6·0	49·8	4·0	54·9

Sex	Frontal Max. Sub. to $N\beta$		Occipital Max. Sub. to λ Op.		Max. Sub. to $N\lambda$		Max. Sub. to G. I.	
	x from N	y	x from λ	y	x from N	y	x from G	y
♂	50·9	24·6	45·2§	26·8§	84·9	68·7	104·5	100·0
♀	50·2	26·7	42·5	26·2	81·9	69·2	100·0	98·4

* These lengths are the means of 29 contours.
† ,, ,, ,, 30 ,,
‡ ,, ,, ,, 9 ,,
§ ,, ,, ,, 28 ,,

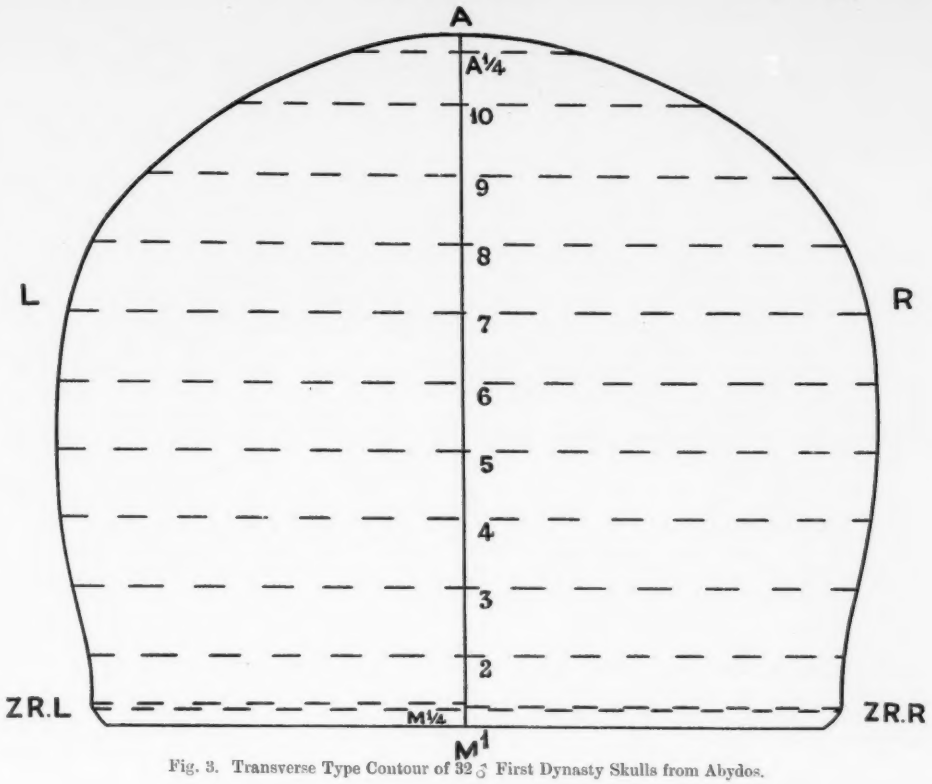


Fig. 3. Transverse Type Contour of 32 ♂ First Dynasty Skulls from Abydos.

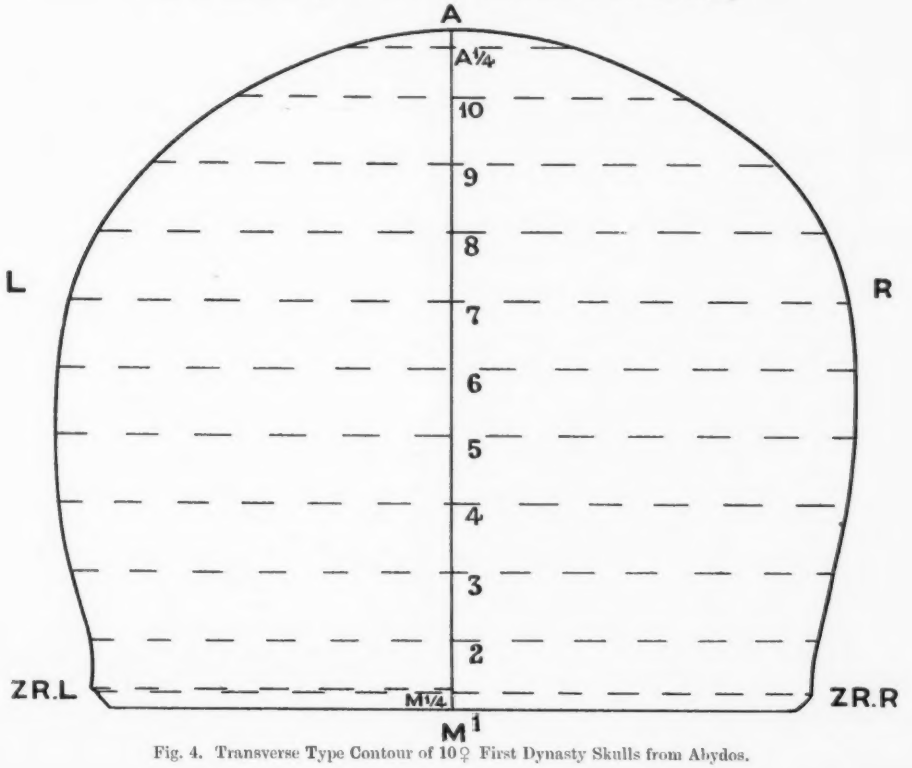


Fig. 4. Transverse Type Contour of 10 ♀ First Dynasty Skulls from Abydos.

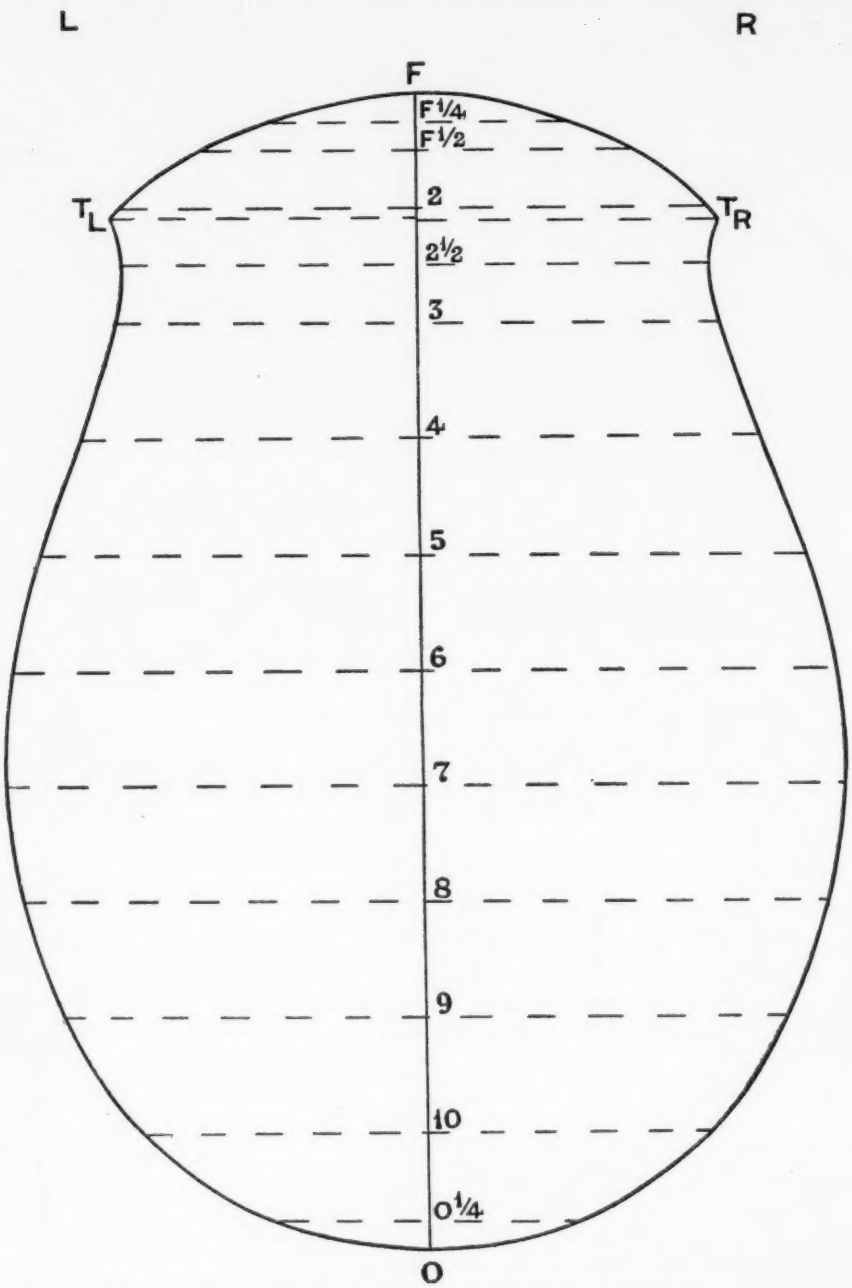


Fig. 5. Horizontal Type Contour of 32 ♂ First Dynasty Skulls from Abydos.

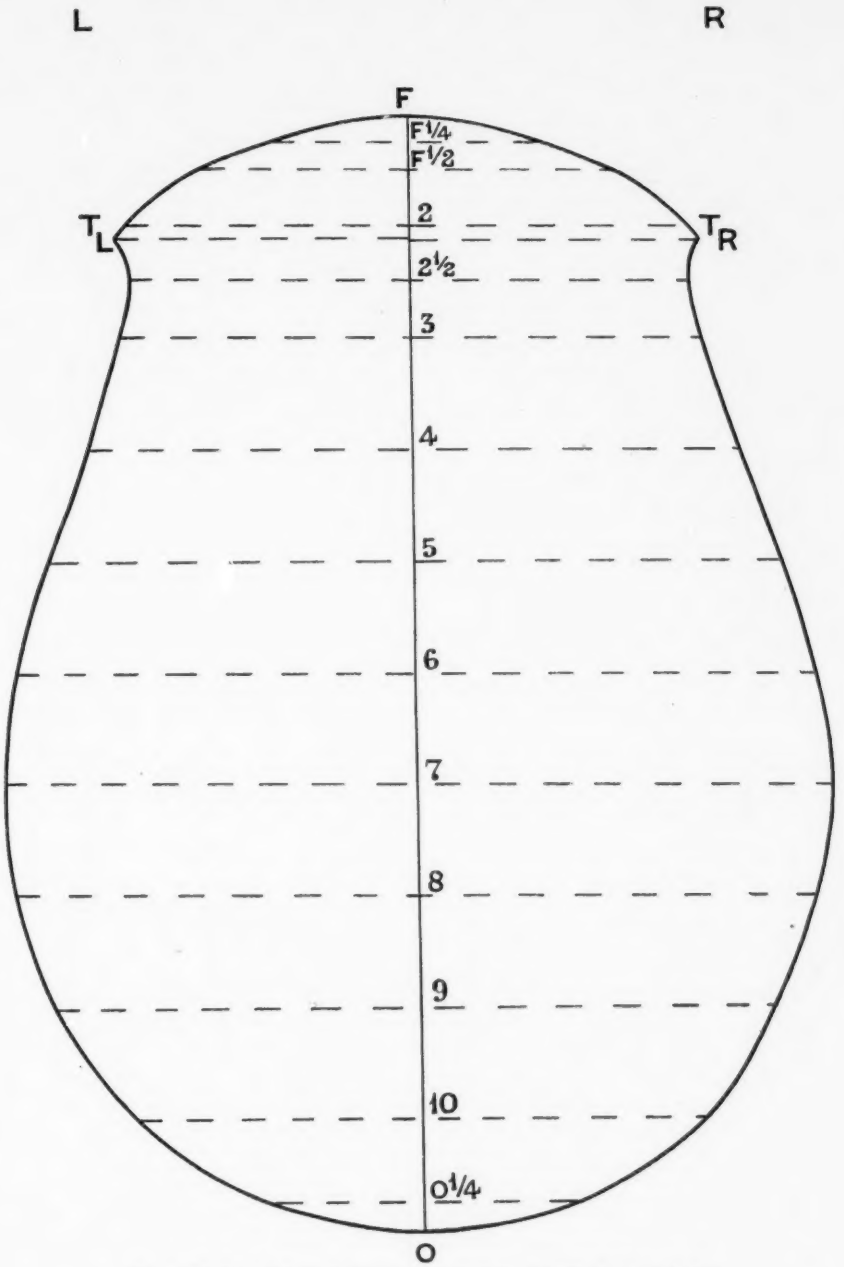


Fig. 6. Horizontal Type Contour of 10 ♀ First Dynasty Skulls from Abydos.

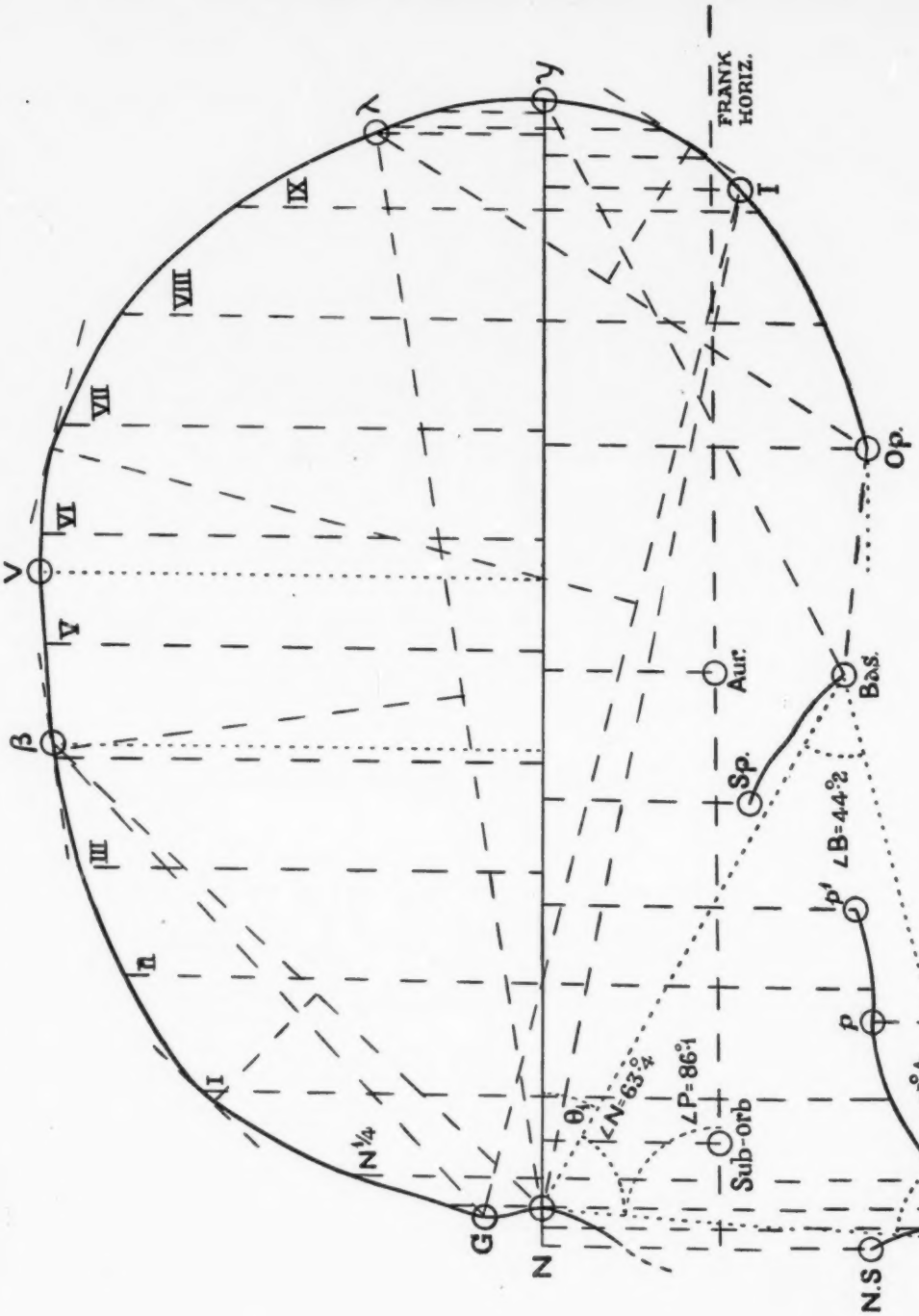


Fig. 7. Sagittal Type Contour of 32 ♂ First Dynasty Skulls from Abydos.

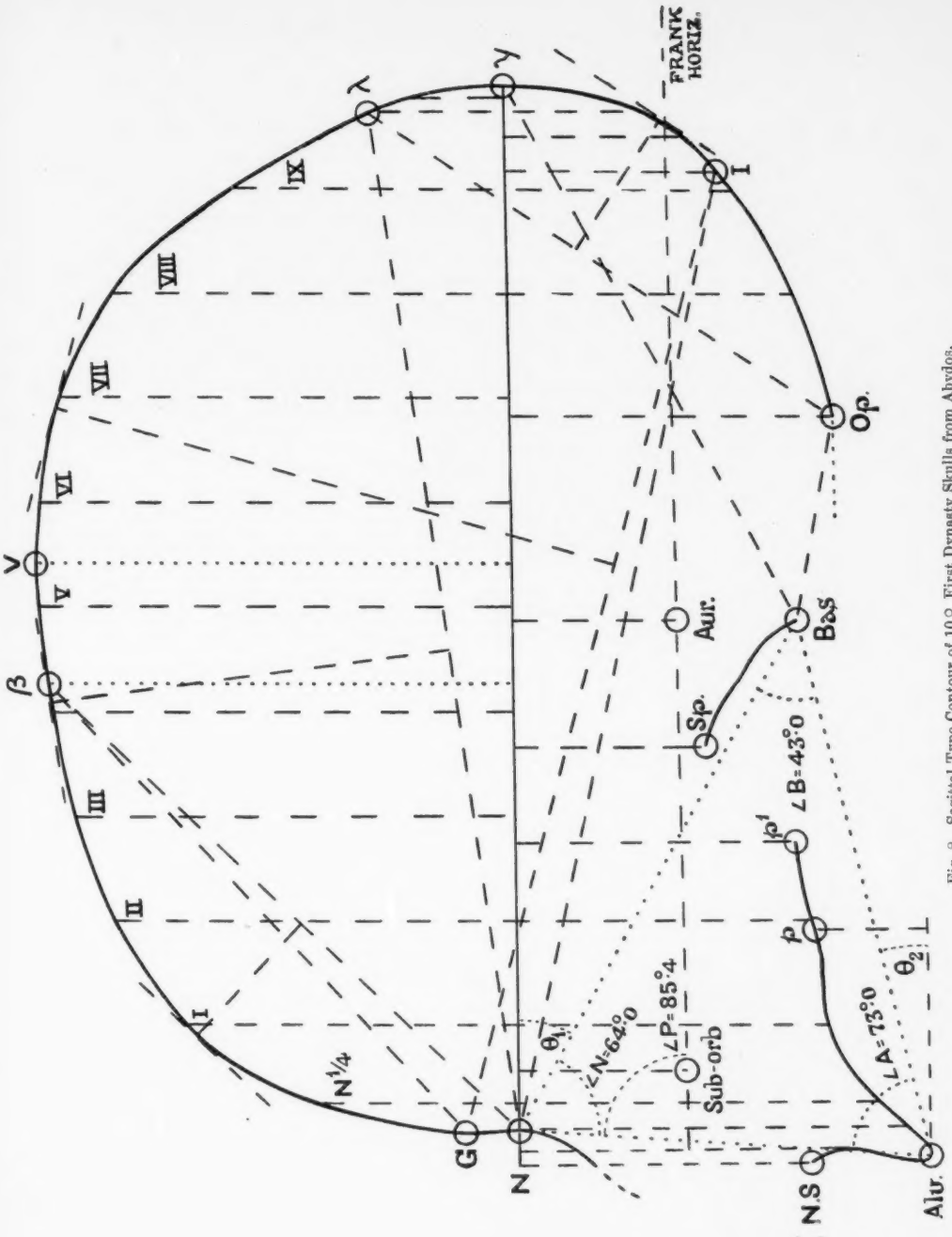


Fig. 8. Sagittal Type Contour of 10 ♀ First Dynasty Skulls from Abydos.

Fig. 7. Sagittal Type Contour of 32 ♂ First Dynasty Skulls from Abydos.

REFERENCES.

(1) Morton, S. G. Observations on Egyptian Ethnology, derived from Anatomy, History, and the Monuments. *Transactions of the American Philosophical Society*, Vol. IX. Philadelphia. 1846. (Paper read in 1842 and 1843.)

(2) Newberry, P. E. *Egypt as a Field for Anthropological Research.* Presidential Address to Section H (Anthropology). *Report of the Britis' Association for the Advancement of Science*, 1923.

(3) Flinders Petrie, W. M. (with Quibell, J. E.). *Naqada and Ballas.* 1896.

(4) —— *Diospolis Parva.* "The Cemeteries of Abadiyeh and Hu." 1901.

(5) —— *Gizeh and Rifeh.* 1907.

(6) Fawcett, C. D. A Second Study of the Variation and Correlation of the Human Skull with special reference to the Naqada Crania. *Biometrika*, Vol. I. pp. 408—467. 1902.

(7) Crewdson Benington, R. and Pearson, Karl. A Study of the Negro Skull with Special Reference to the Congo and Gaboon Crania. *Ibid.* Vol. VIII. pp. 292—339. 1912.

(8) Smith, H. D. A Study of Pygmy Crania, based on skulls found in Egypt. *Ibid.* Vol. VIII. pp. 262—266. 1912.

(9) Warren, E. An Investigation on the Variability of the Human Skeleton: with special reference to the Naqada Race. *Phil. Trans.* Series B, Vol. 189, pp. 135—227. 1897.

(10) Morgan, J. de. *Recherches sur les Origines de l'Égypte.* "L'Age de la Pierre et les Métaux." 1896.

(11) —— *Recherches sur les Origines de l'Égypte.* "Ethnographie Préhistorique et Tombeau Royal de Négadah." 1897.

(12) Randall-MacIver, D. *The Earliest Inhabitants of Abydos.* 1901.

(13) Thomson, A. and Randall-MacIver, D. *The Ancient Races of the Thebaid.* 1905.

(14) Chantre, E. Étude Craniologique sur la population prépharaonique de la Haute Égypte. *Bulletin de la Société d'Anthropologie de Lyon.* Tome 18, pp. 187—192. 1899.

(15) —— *Recherches Anthropologiques dans l'Afrique Orientale.* "Égypte." 1904.

(16) Quatrefages, A. de and Hamy, E. T. *Crania Ethnica.* "Les Cranes des Races Humaines." 1882.

(17) Pruner-Bey. Recherches sur l'Origine de l'ancienne Race Égyptienne. *Mémoires de la Société d'Anthropologie de Paris.* T. 1, pp. 399—433. 1861—3. (Paper read in 1861.)

(18) Biasutti, R. "Crania Aegyptiaca." Esame di 42 Crani di Egiziani Antichi. *Archivio per l'Antropologia e la Etnologia.* Vol. 35, pp. 323—362. Firenze. 1905.

(19) Die Anthropologischen Sammlungen Deutschlands. (Dr Emil Schmidt's catalogue of his own collection at Leipzig was published with *Archiv für Anthropologie.* Bd. XVII. 1887.)

(20) Toldt, C. Anthropologische Untersuchung der Menschlichen Ueberreste aus den Altägyptischen Gräberfeldern von El-Kubanieh. *Denkschriften der Akademie der Wissenschaften in Wien.* Mathematisch-naturwissenschaftliche Klasse. Bd. 96, pp. 593—672. 1919.

(21) Stahr, Hermann. *Die Rassenfrage in Antiken Aegypten.* "Kraniologische Untersuchungen an Mumienköpfen aus Theben." Leipzig. 1907.

(22) Oetteking, B. Kraniologische Studien an Altägypten. *Archiv für Anthropologie.* Bd. XXXVI. pp. 1—90. 1909.

(23) Schultz, A. Anthropologische Untersuchungen an der Schädelbasis. *Ibid.* Bd. XLIV. pp. 1—103. 1918.

(24) Widenmann. Untersuchung von 30 Dschaggaschädeln. *Ibid.* Bd. XXV. pp. 361—396. 1898.

(25) Drontschilow, Krum. Metrische Studien an 93 Schädeln aus Kamerun. *Ibid.* Bd. XL. pp. 161—183. 1913.

(26) Sergi, Sergio. *Crania Habessinica. Contributo all' Antropologia dell' Africa Orientale.* 1912.

APPENDIX.

			Lengths									Arcs				
No.	*C*	*W*	*F*	*L'*	*L*	*B*	*B'*	*H*	*H'*	*OH*	*LB*	*Q*	*Q'*	*S*	S_1	S_2
Males																
131	1256	596	174·8	177·5	176·3	136·5	94·9	131·3	131·6	112·7	102·0	302·0	302·5	357·0	120·5	131·5
132	—	535?	176·9	—	180·7?	137·1	95·3	—	118·7?	—	101·1	—	—	355·5	117·5	132·0
146	1438	954·5	191·5	193·2	191·5	144·4	99·0	145·0	143·6	120·2	108·1	322·0	324·0	394·5	134·5	137·0
153 A	1639	860	196·9	201·0	200·4	134·8	94·0	145·8	146·4	120·3	110·2	322·0	322·0	396·0	128·0	144·0
205	1276	660	181·1	186·8	185·5	136·4	89·2	127·8	127·0	109·8	100·6	295·5	295·5	363·0	126·0	128·0
229	1399	558·5	184·8	187·5	187·2	136·0	93·0	125·5	124·0	112·0	101·4	309·7	310·0	367·0	125·5	124·5
254	1278	824	183·2	184·0	183·8	132·6	99·1	129·0	127·6	113·9	93·9	303·5	310·0	379·0	129·5	135·0
396	1478	589·5	187·9	191·0	190·1	147·0	93·2	134·9	133·0	113·7	103·4	313·0	313·0	380·0	121·0	121·0
413	1297	673	176·0?	179·1	178·9?	132·6	84·1	142·0	144·1?	117·1	100·2	310·0	312·0	369·0	124·0	132·0
426	1410	699	180·2	181·5	181·0	140·4	88·5	134·0	133·5	113·3	100·0	308·5	308·5	376·0	126·0	130·0
428	—	—	184·3?	—	185·3?	139·0	94·8	—	121·4	—	—	—	—	—	—	—
429	1323	616	181·2	185·8	184·3	136·0	88·5	133·0	132·2	112·2	104·0	304·0	304·0	366·0	131·0	122·0
432	1441	591	184·2	186·0	185·9	140·7	93·0	137·0	136·2	115·0	100·0	309·5	309·5	381·0	131·0	125·0
433	1350	762	184·2	187·2	186·5	139·0	97·0	133·7	133·8	115·0	105·8	313·5	315·5	372·0	130·0	134·0
436	—	608	182·1	185·2	185·0	135·9	95·0	137·6	135·1	112·7	101·4	301·5	303·0	373·0	123·0	130·0
442	1366	676	180·1?	183·0	182·0?	132·1	94·5	137·0	136·7?	114·9	102·1	304·2	305·0	370·0?	124·0?	136·0
446	1438	651·5	184·1	185·0	185·0	138·6	88·4	138·0	137·5	120·9	99·0	316·3	320·0	381·0	131·0	127·0
447	1203	800	184·3	191·5	190·9	132·0	96·0	134·0	131·1	111·2	108·8	295·7	296·0	366·0	121·0	129·0
449	1274	553·5	183·1	189·7	189·0	133·0	89·5	134·0	129·7	110·5	105·5	300·5	301·0	365·0	115·0	129·0
452	—	616?	177·0	183·6	183·1	126·0	84·1	—	—	106·0	—	284·2	284·5	—	110·0	130·0
454	—	745·5?	181·8	—	182·6	139·0	—	—	133·4	—	101·9	—	—	—	125·0	136·5
461	—	680·5?	180·4	185·0	184·4	136·0	78·0	133·7	132·7	113·8	94·8	311·5	312·0	—	124·5	137·5
484	1292	692	180·1	183·2	183·0	139·0	99·2	135·6	134·9	114·9	104·1	310·0	310·0	364·0	133·0	118·0
510	1362	613·5	182·1	185·2	185·0	136·7	91·0	128·1	128·3	107·8	100·0	296·2	297·5	373·0	123·5	127·0
511	1424	626·5?	187·1?	187·7	188·1?	136·0	100·0	136·1	135·5	114·1	102·0	306·2	305·5	—	—	124·0
530	—	—	175·0?	—	177·5?	134·0?	—	—	122·1	—	—	—	—	352·0	118·0	122·0
533	—	—	—	—	—	132·6	97·0	—	137·8?	—	—	—	—	—	—	—
543	—	—	190·0?	—	197·6?	141·7	100·0	—	—	—	—	—	—	—	—	—
602	—	664	180·4	185·8	184·1	137·6	94·0	131·7	131·6	111·6	102·9	299·5	299·0	361·0	120·0	127·0
603	—	552?	175·0	176·8	177·2	133·0	91·1	129·0	129·2	103·0	100·0	285·0	284·0	349·0	114·0	122·0
675	1332	707	177·0	179·2	178·8	136·0	95·8	128·0	127·5	107·6	93·2	293·5	294·0	362·0	120·0	126·0
740	1470	717	191·3	195·0	194·9	139·7	96·6	133·6	133·9	105·0	108·0	309·0	309·0	380·0	131·5	132·5
745	—	616	180·0	184·8	184·5	135·0	93·0	128·2	126·6	104·0	102·8	307·0	308·0	361·0	118·0	134·0
746	—	—	184·2	188·2	187·2	128·5?	92·0	139·0	139·0	116·2	110·9	309·5?	—	365·0	121·5	130·5
756	1352	626	176·0	177·0	176·6	136·0	92·8	131·4	130·5	110·8	100·0	299·5	299·5	358·5	121·0	131·0
763	1390	651	182·0	184·5	184·1	135·0	90·0	132·1	130·2	113·7	98·4	304·0	304·0	364·0	121·0	131·0
776	1254	684	179·0?	180·0	179·5?	135·4	92·8	126·4	127·2	111·8	99·5?	304·0	306·0	—	—	123·5
Negro																
443	1375	—	181·5	184·2	183·1	142·0	103·8	132·8	131·1	111·9	104·1	309·0	310·0	369·0	122·5	134·0
Females																
135	1138	637	175·0	177·7	175·2	130·1	88·2	127·3	127·3	106·2	99·7	289·2	291·0	350·0	123·5	115·0
153B	1254	578	174·5	176·6	175·5	134·2	88·8	127·7	126·1	111·9	95·4	298·5	299·0	365·5	120·5	132·5
302	1179	689	176·5	176·3	176·0	130·3	89·0	134·1	134·0	110·1	101·4	292·5	294·0	355·5	122·0	127·0
412	1175	608	168·1	172·2	178·0	129·1	85·7	127·2	123·4	109·5	99·1	294·0	296·0	346·0	116·0	116·0
437	—	535?	178·0	180·0	179·5	136·0	89·1	126·6	126·3	108·0	97·0	295·0	296·0	359·0?	127·0	121·0
444	—	644	176·7	178·0	177·3	136·0	92·4	132·5	131·9	117·5	102·3	296·0	297·5	356·0	131·5	118·0
445	—	—	185·1	187·4	186·6	140·2	93·8	136·8	136·0	107·8	100·6	316·5	—	377·0	133·0	140·0
470	1331	617	185·0?	187·1?	186·2?	134·4	89·0	131·1	129·7	112·9	102·0	303·0	301·0	—	—	132·0
538	—	—	192·0?	190·5?	191·4?	135·9	96·3	—	—	116·0	—	311·0	311·5	—	144·0?	—
700	1230	469	174·0	176·3	174·4	128·0	88·2	125·1	123·4	108·0	97·4	288·5	290·0	353·5	121·0	117·5

Arcs					Face				Nose							Orbita				
S_1	S_2	S_3	S_3'	U	PH	$G'H$	GB	J	$\frac{NH}{R}$	$\frac{NH}{L}$	NB	DS	DC	DA	SS	SC	O_1R	O_1L	O_2R	O_2L	$O_1'R$
120·5	131·5	105·0	88·0	496·0	21·0	72·5	93·5	128·1	51·0	51·7	23·6	13·13	23·1	37·4	5·56	9·6	42·0	40·5	30·6	31·8	39·0
117·5	132·0	106·0	90·7?	511·0	18·9	—	94·1	134·6	48·8	48·6	27·4	11·10	22·1	37·5	4·08	14·7	44·5	42·0	30·5	30·2	42·8
134·5	137·0	123·0	103·0	537·0	19·0	76·2	98·3	138·2	57·6	57·0	26·4	14·05	25·9	42·1	5·70	15·2	43·8	43·1	32·9	32·9	41·0
128·0	144·0	124·0	105·0	540·0	21·3	68·6	99·8	138·6	49·0	49·5	27·6	13·15	24·0?	32·0	5·28	13·6	46·0	44·7	31·6	30·0	41·5
126·0	128·0	109·0	89·1	510·0	24·6	84·2	99·8	127·0	59·0	59·9	25·5	11·56	23·0	35·2	4·92	12·0	45·0	41·0	35·1	37·0	41·0
125·5	124·5	117·0	93·4	525·0	16·0	70·0	88·0	131·4	53·8	53·0	26·0	12·65	21·0	32·9	4·30	11·6	44·9	43·5	35·5	34·7	43·0
129·5	135·0	114·5	93·9	512·0	22·2	71·9	99·0	124·0	50·7	49·8	28·5	9·06	25·4?	32·6	3·05	11·8	44·5	42·8	31·5	31·4	40·4
121·0	121·0	138·0	108·0	531·0	20·1	76·0	96·0	125·2	55·1	55·6	26·4?	—	19·4?	—	—	11·0	44·0	46·9	35·7	34·0	—
124·0	132·0	113·0	99·8	490·0	19·9	76·0	91·8	120·7	59·0	59·9	23·3	11·69	20·7	33·5	6·19	11·1	41·8	40·5	34·2	33·0	38·4
126·0	130·0	120·0	97·3	510·0	25·8	78·9	92·1	124·3	53·6	55·8	26·9	9·65	22·0	30·0	3·43	10·3	41·9	41·3	36·4	35·2	39·0
—	—	—	85·0?	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
131·0	122·0	113·0	89·2	509·0	22·8	80·0	103·8	133·0	58·9	58·9	25·8	14·04	25·6	39·8	5·00	10·6	41·5	43·3	32·9	33·7	38·9
131·0	125·0	125·0	98·6	519·0	22·1	75·7	97·7	138·0?	55·1	57·1	26·6	12·16	20·7	35·0	5·41	10·9	42·0	41·0	35·1	36·9	40·3
130·0	134·0	108·0	88·7	524·0	20·7?	78·8	98·1	128·5	58·5	58·1	25·0	12·14	24·0	34·0	3·50	12·0	40·0	40·5	31·6	32·2	38·0
123·0	130·0	120·0	101·8	521·0	16·8	73·0	94·6	128·9	57·5	57·0	25·1	12·10	22·9	36·0	4·63	8·9	42·2	43·0	33·5	33·0	40·0
124·0?	136·0?	110·0	93·9	504·0	22·2	79·5	96·1	133·0	57·9	57·1	26·0	14·59	23·0	40·1	5·30	11·8	42·9	42·9	34·9	35·0	39·0
131·0	127·0	123·0	102·8	514·0	23·0	71·0	111·0	126·1	50·4	49·5	26·0	14·17	23·6	39·7	4·53	12·4	41·9	41·6	33·0	34·0	39·2
121·0	129·0	116·0	93·8	517·0	18·4	72·9	99·2	130·5	53·2	53·9	27·3	11·48	26·8	35·2	4·20	9·8	44·0	43·6	34·0	32·5	39·5
115·0	129·0	121·0	99·2	515·0	17·0	71·3?	97·8	133·0	55·1	54·8	27·0	13·73	24·6	37·0	6·66	10·9	42·9	41·5	30·9	31·0	39·4
110·0	130·0	—	—	495·0	20·9	74·5	95·8	127·1	54·2	53·9	24·7	13·75	22·0	37·0	3·70	11·6	42·1	42·9	35·0	38·3	40·5
125·0	136·5?	—	—	—	21·8?	—	98·8	—	52·2	52·2	26·5	—	—	—	—	12·5?	43·1	—	31·0	—	39·5
124·5	137·5	—	—	—	27·5?	81·0?	97·4	—	56·1	54·7	24·4	11·20?	22·0	35·0	3·98?	9·6	43·0	42·5	33·0	37·5	41·4
133·0	118·0	112·0	93·0	518·0	20·9	72·9	104·0	130·2	51·8	51·8	24·6	11·43?	26·0	38·8	2·69?	9·0	41·1	42·2	32·0	32·9	39·6
123·5	127·0	122·5	101·0	509·0	19·7	74·8	94·1	125·0	56·0	57·1	25·9	12·57	24·0	36·1	4·49	11·0	44·8	43·2	36·1	36·3	42·4
—	124·0?	125·0	100·9	524·0	21·0	76·7	96·5	130·0	—	55·6	28·7	8·90	24·9	32·6	2·82	13·4	44·1	42·0	32·4	35·1	40·8
118·0	122·0	112·0	94·8	—	22·8	—	92·2	—	—	—	22·8	—	—	—	—	—	—	—	—	—	—
—	—	121·5	100·1	—	—	—	—	—	—	—	—	—	21·0?	—	—	8·6	—	—	—	—	—
—	—	—	—	—	—	—	—	—	—	—	—	—	19·7?	—	—	—	—	—	—	—	—
120·0	127·0	114·0	94·2	518·0	23·9	75·0	98·9	125·6	52·1	50·9	22·2	11·90	22·5	36·1	5·16	8·3	41·9	41·4	32·9	34·1	41·3
114·0	122·0	113·0	94·0	492·0	23·9	79·2	97·0	127·2	55·8	56·0	23·7	11·53	22·1	34·5	5·09	9·5	42·0	43·1	36·6	36·6	40·0
120·0	126·0	116·0	95·2	506·0	20·9	74·2	102·7	132·9	55·2	56·1	24·1	14·07	22·8	37·5	4·65	10·2	41·0	41·1	34·9	34·9	38·2
131·5	132·5	116·0	95·0	537·0	23·6?	75·0	107·0	137·6	51·9	51·4	28·0	13·01?	24·3	37·0?	3·78	10·9	43·0	43·2	31·1	33·3	40·8
118·0	134·0	109·0	92·7	505·0	20·1	70·6	92·9	127·9	51·0	49·8	23·3	10·73	24·3	35·0	4·33	10·7	41·9	41·0	31·0	31·0	38·7
121·5	130·5	113·0	96·1	—	19·1?	74·0	—	125·5	54·1	—	—	—	17·9?	—	5·41	9·7	42·5	—	34·8	—	40·0
121·0	131·0	106·5	93·1	501·0	22·8?	71·0	94·1	132·9	47·9	49·0	23·3	11·70	21·7	34·0	4·24	10·1	41·6	42·1	33·0	33·7	39·7
121·0	131·0	112·0	93·0	511·0	24·8	76·5	94·2	128·0	52·8	52·0	24·8	12·10	20·0	34·5	5·34	9·8	42·5	42·2	35·1	33·9	40·0
—	123·5?	113·0	92·5	505·0	18·2	63·9	92·9	127·0	48·8	48·8	25·2	—	22·9	—	—	10·0	40·0	40·2	33·0	33·5	38·0
122·5	134·0	113·5	96·4	513·0	21·8	76·5	104·0	142·1	68·7	68·1	29·0	14·38	29·0	46·1	2·59	13·0	45·0	44·0	36·0	34·3	41·0
123·5	115·0	111·0	91·2	490·0	27·7	76·7	98·8	123·9	53·5	53·0	24·4	11·11	21·7	33·0	3·26	10·2	40·9	39·6	32·2	32·1	38·7
120·5	132·5	112·0	94·3	488·5	20·0	67·2	93·9	118·3	47·6	47·6	25·4	8·97	24·9	31·5	2·48	8·0	40·6	40·6	33·0	33·6	39·1
122·0	127·0	107·0	89·3?	488·5	17·5	65·7	89·1	114·8	51·1	52·0	23·9	12·59	20·0	33·6	4·07	11·0	40·2	39·0	33·1	33·1	37·1
116·0	116·0	114·0	95·6	481·0	22·0	68·6	92·5?	126·5	48·1	49·2	24·0	10·70	20·3	33·0	3·73	8·6	41·2	41·2	32·2	30·6	39·0
127·0	121·0	110·0?	88·3?	502·0	17·0?	73·9	92·1	122·7	54·9	54·8	25·3	10·46	20·3	32·0	3·90	8·5	40·1	39·0	31·0	31·1	37·6
131·5	118·0	106·0	92·9	502·0	16·2	70·2	98·0	129·3	53·4	53·4	24·8	12·30	21·0	32·5	2·94	8·1	40·2	40·2	35·6	35·1	37·7
133·0	140·0	104·0	89·4	—	20·6?	76·7	—	—	55·1	55·2	24·9	8·73	21·1	30·5	3·14	8·1	41·9	—	34·5	—	41·8
—	132·0?	114·0	92·0	506·0	17·8	67·0	98·9	126·0	49·0	47·9	23·2	10·85	18·5	34·5	4·62	10·9	41·1	40·0	33·2	33·4	39·0
144·0?	—	125·0?	97·0	521·5	21·1	68·6	96·1	119·0	48·8	48·3	28·0	11·90	27·0	36·5	3·22	12·5	41·2	41·4	33·1	33·1	38·5
121·0	117·5	115·5?	95·4	492·5	24·1	73·0	85·6	116·6	49·7	49·8	21·3	12·55	20·4	33·5	5·27	9·2	40·8	39·0	32·6	33·0	36·5

Orbits				Palate and Profile						Foramen									
O_2R	O_2L	$O_1'R$	EOW	G_1	G_1'	G_2	GL	EB	EH	fml	fmb	$Oc.I$	B/L'	H/L'	B/L	H/L	B/H	H'/L	B/H'	$\frac{(B-H')}{L}$
30·6	31·8	39·0	96·5	50·9	46·5	43·9	93·0	44·0	15·2	34·2	30·0	60·5	76·9	74·0	77·4	74·5	104·0	74·6	103·7	2·8
30·5	30·2	42·8	102·1	—	—	—	—	—	—	35·0	29·3	62·8	—	—	75·9	—	—	65·7?	115·5?	10·2?
32·9	32·9	41·0	98·6	57·1	49·6	—	99·0	—	—	35·0	31·8	60·5	74·7	75·1	75·4	75·7	99·6	75·0	100·6	0·4
31·6	30·0	41·5	104·4	55·0	51·0	42·3	105·0	42·8	12·0	41·1	35·1	61·6	67·1	72·5	67·3	72·8	92·5	73·1	92·1	−5·8
35·1	37·0	41·0	95·5	55·9?	50·3	43·2	101·7	43·2	14·9	40·3	—	58·4	73·0	68·4	73·5	68·9	106·7	68·5	107·4	5·1
35·5	34·7	43·0	101·1	57·0	49·2	40·2	96·2	40·5	10·5	37·1	29·1	56·9	72·5	66·9	72·6	67·0	108·4	66·2	109·7	4·4
31·5	31·4	40·4	100·2	51·1	47·0	41·7	92·0	41·8	16·8	38·7	30·4	58·7	72·1	70·1	72·1	70·2	102·8	69·4	103·9	2·7
35·7	34·0	—	97·4	49·3?	48·1	41·1	90·1	41·0	16·2	36·6	31·4	55·9	77·0	70·6	77·3	71·0	109·0	70·0	110·5	7·4
34·2	33·0	38·4	93·4	53·2	48·2	40·2	89·2	40·5	23·6?	36·9	30·2	67·7	74·0	79·3	74·1?	79·4	93·4	80·5?	92·0?	−6·4?
36·4	35·2	39·0	94·0	50·1	46·0	42·2	90·1	42·5	16·9	35·2	26·0	57·9	77·4	73·8	77·6	74·0	104·8	73·8	105·2	3·8
—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	75·0?	—	—	65·5	114·5	9·5?
32·9	33·7	38·9	96·6	54·1	51·5	46·9	95·5	47·0	8·5	35·0	27·7	56·3	73·2	71·6	73·8	72·2	102·3	71·7	102·9	2·1
35·1	36·9	40·3	96·9	51·7	47·3	—	91·0	—	—	34·0	29·0	56·3	75·6	73·7	75·7	73·7	102·7	73·3	103·3	2·4
31·6	32·2	38·0	99·2	55·9	52·8	40·8	103·1	41·0	18·9	32·0	26·9	58·8	74·3	71·4	74·5	71·7	104·0	71·7	103·9	2·8
33·5	33·0	40·0	100·1	57·2	50·4	42·3	97·0	42·5	14·3	37·3	36·8	61·8	73·3	74·3	73·4	74·4	98·6	73·0	100·4	0·3
34·9	35·0	39·0	99·0	59·1	50·0	46·0	98·7	46·0	14·3	40·0	33·8?	62·5	72·2	74·9	72·6	75·3	96·4	75·1?	96·6?	−2·5?
33·0	34·0	39·2	96·3	53·0	47·4	40·8	95·9	40·9	12·8	35·0	30·2	60·3	74·9	74·6	74·9	74·6	100·4	74·3	100·8	0·6
34·0	32·5	39·5	100·2	55·0	50·6	43·3	102·5	43·0	7·3	37·2	26·1	57·7	68·9	70·0	69·2	70·2	98·5	68·7	100·7	0·5
30·9	31·0	39·4	98·6	57·1	51·9	44·5	99·1?	44·0	12·6	35·6	27·3	58·7	70·1	70·6	70·4	70·9	99·3	68·6	102·5	1·7
35·0	38·3	40·5	93·4	55·0	47·5	39·5	—	39·5	15·1	—	—	—	68·6	—	68·8	—	—	—	—	—
31·0	—	39·5	—	59·0	53·1	—	—	—	—	37·5	31·6	—	—	—	76·1	—	—	73·1	104·2	3·1
33·0	37·5	41·4	103·0	54·9?	51·0	46·0?	93·5?	46·0	13·4	—	—	—	73·5	72·3	73·8	72·5	101·7	72·0	102·5	1·8
32·0	32·9	39·6	101·1	54·1	47·1	42·2	96·5	42·0	13·6	38·3	28·2	59·7	75·9	74·0	76·0	74·1	102·5	73·7	103·0	2·2
36·1	36·3	42·4	98·9	58·9	54·1	42·6	101·6	42·7	17·6	34·9	29·9	59·1	73·8	69·2	73·9	69·3	106·7	69·4	106·6	4·5
32·4	35·1	40·8	101·3	53·5	47·8	45·1	91·2	45·5	15·7	35·5	26·9	57·6	72·5	72·5	72·3?	72·4	99·9	72·0?	100·4	0·3?
—	—	—	—	55·6	50·1	42·9	—	42·5	11·3	34·5	28·0	61·6	—	—	75·5	—	—	68·8?	109·7	6·7?
—	—	—	99·0	—	—	—	—	—	—	35·5?	28·0	59·0	—	—	—	—	—	—	96·2?	—
—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	71·7	—	—	—	—	—
32·9	34·1	41·3	98·9	57·8	49·0	43·3	98·0	43·5	17·0	39·6	31·0	59·3	74·1	70·9	74·7	71·5	104·5	71·5	104·6	3·3
36·6	36·6	40·0	99·4	54·2?	51·0	45·9	94·0	46·0	19·1	36·5	31·0	59·9	75·2	73·0	75·1	72·8	103·1	72·9	102·9	2·2
34·9	34·9	38·2	99·0	48·9?	44·0	42·2	89·0	42·5	12·0	41·3	32·3	58·8	75·9	71·4	76·1	71·6	106·2	71·3	106·7	4·8
31·1	33·3	40·8	103·2	61·9?	55·1	46·2	107·0	46·5	16·9	40·3	31·1	58·6	71·6	68·5	71·7	68·5	104·6	68·7	104·3	3·0
31·0	31·0	38·7	97·8	52·9?	49·1	41·1	98·5	41·0	16·2	36·9	29·2	62·1	73·1	69·4	73·2	69·5	105·3	68·6	106·6	4·6
34·8	—	40·0?	—	—	—	—	96·6?	—	—	38·8	28·9	62·1	68·3?	73·9	68·6?	74·2	92·4	74·3	92·4?	5·6
33·0	33·7	39·7	95·9	51·1?	48·9	42·4	93·5?	41·2	16·8	38·9	32·6	65·8	76·8	74·2	77·0	74·4	103·5	73·9	104·2	3·1
35·1	33·9	40·0	94·6	51·9	45·6	41·4	95·0	41·6	16·0	41·1	33·9	59·7	73·2	71·6	73·3	71·8	102·2	70·7	103·7	2·6
33·0	33·5	38·0	93·7	56·8	52·1	40·5	92·0	40·2	11·5	35·2	28·0	59·2	75·2	70·2	75·4	70·4	107·1	70·8	106·4	4·6?
36·0	34·3	41·0	107·7	54·1	49·5	44·7	104·1	44·7	15·9	34·6	35·1	61·9	77·1	72·1	77·6	72·5	106·9	71·6	108·3	6·0
32·2	32·1	38·7	94·4	58·3	51·8	41·5	97·0	41·6	16·4	35·4	28·0	58·8	73·2	71·6	74·3	72·7	102·2	72·7	102·2	1·6
33·0	33·6	39·1	97·0	53·0?	47·7	—	93·8	—	—	34·1	28·6	61·0	76·0	72·3	76·5	72·8	105·1	71·9	106·4	4·6
33·1	33·1	37·1	89·9	52·0	47·1	39·0	95·1	39·0	11·5	38·9	27·5	60·2	73·9	76·1	74·0	76·2	97·2	76·1	97·2	−2·1
32·2	30·6	39·0	94·1	49·0	45·1	—	91·2	—	—	31·5	26·8	60·6	75·0	73·9	75·6	74·5	101·5	72·2	104·6	3·4
31·0	31·1	37·6	91·3	53·2?	48·4	40·5	92·0	40·5	14·1	35·9	30·0	56·8	75·6	70·3	75·8	70·5	107·4	70·4	107·7	5·4
35·6	35·1	37·7	93·1	54·9?	49·0	41·2	95·0	41·0	11·8	37·0	32·1	66·3	76·4	74·4	76·7	74·7	102·6	74·4	103·1	2·3
34·5	—	41·8	—	60·4	52·0	40·9	92·3	40·8	13·2	38·1	29·0	63·4	74·8	73·0	75·1	73·3	102·5	72·9	103·1	2·2
33·2	33·4	39·0	91·2	53·0	47·0	40·2	97·6	40·0	13·0	38·0	28·1	57·6	71·8	70·1	72·2	70·4	102·5	69·7	103·6	2·5
33·1	33·1	38·5	95·2	51·0	47·2	41·1	—	40·8	10·0	—	—	55·6	71·3	—	71·0	—	—	—	—	—
32·6	33·0	36·5	93·0	51·8?	47·0	40·4	90·7	40·0	15·0	34·1	28·9	59·3	72·6	71·0	73·4	71·7	102·3	70·8	103·7	2·6

men | Indices

fmb	*Oc.I*	*B/L′*	*H/L′*	*B/L*	*H/L*	*B/H*	*H′/L*	*B/H′*	$\frac{(B-H')}{L}$	*G′H/GB*	*NB/NH R*	*NB/NH L*	O_3/O_1 *R*	O_3/O_1 *L*	O_2/O_1' *R*	G_2/G_1	*EH/EB*	*fmb/fml*	*DS/DC*
30·0	60·5	76·9	74·0	77·4	74·5	104·0	74·6	103·7	2·8	77·5	46·3	45·6	72·9	78·5	78·5	86·2	34·5	87·7	56·8
29·3	62·8	—	—	75·9	—	—	65·7?	115·5?	10·2?	—	56·1	56·4	68·5	71·9	71·3	—	—	83·7	50·2
31·8	60·5	74·7	75·1	75·4	75·7	99·6	75·0	100·6	0·4	77·5	45·8	46·3	75·1	76·3	80·2	—	—	90·9	54·2
35·1	61·6	67·1	72·5	67·3	72·8	92·5	73·1	92·1	− 5·8	68·7	56·3	55·8	68·7	67·1	76·1	76·9	28·0	85·4	54·8?
—	58·4	73·0	68·4	73·5	68·9	106·7	68·5	107·4	5·1	84·3	43·2	42·6	78·0	90·2	85·6	77·3	34·5	—	50·3
29·1	56·9	72·5	66·9	72·6	67·0	108·4	66·2	109·7	4·4	79·5	48·3	49·1	79·1	79·8	82·6	70·5	25·9	78·4	60·2
30·4	58·7	72·1	70·1	72·1	70·2	102·8	69·4	103·9	2·7	72·6	55·1	57·2	70·8	73·4	78·0	81·6	40·2	78·6	35·7?
31·4	55·9	77·0	70·6	77·3	71·0	109·0	70·0	110·5	7·4	79·2	47·9	47·5	81·1	72·8	—	83·4	39·5	85·8	—
30·2	67·7	74·0	79·3	74·1?	79·4	93·4	80·5?	92·0?	− 6·4?	82·8	39·5	38·9	81·8	81·5	80·1	75·6	58·3	81·8	56·5
26·0	57·9	77·4	73·8	77·6	74·0	104·8	73·8	105·2	3·8	85·7	50·2	48·2	86·9	85·2	93·3	84·2	39·8	73·9	43·9
—	—	—	—	75·0?	—	—	65·5	114·5	9·5?	—	—	—	—	—	—	—	—	—	—
27·7	56·3	73·2	71·6	73·8	72·2	102·3	71·7	102·9	2·1	77·1	43·8	43·8	79·2	77·8	86·6	86·7	18·1	79·1	54·8
29·0	56·3	75·6	73·7	75·7	73·7	102·7	73·3	103·3	2·4	77·5	48·3	46·6	83·6	90·0	87·1	—	—	85·3	58·7
26·9	58·8	74·3	71·4	74·5	71·7	104·0	71·7	103·9	2·8	80·3	42·7	43·0	79·9	79·5	83·2	73·0	46·1	84·1	50·6
36·8	61·8	73·3	74·3	73·4	74·4	98·6	73·0	100·4	0·3	77·2	43·6	44·0	79·4	76·7	83·8	74·0	33·6	98·7	52·8
33·8?	62·5	72·2	74·9	72·6	75·3	96·4	75·1?	96·6?	− 2·5?	82·7	44·9	45·5	81·3	82·6	89·5	77·8	31·1	84·5?	63·4
30·2	60·3	74·9	74·6	74·9	74·6	100·4	74·3	100·8	0·6	64·0	51·6	52·5	78·8	81·7	84·2	77·0	31·3	86·2	60·0
26·1	57·7	68·9	70·0	69·2	70·2	98·5	68·7	100·7	0·5	73·5	51·3	50·7	77·3	74·5	86·1	78·7	17·0	70·2	42·8
27·3	58·7	70·1	70·6	70·4	70·9	99·3	68·6	102·5	1·7	72·4?	49·0	49·3	72·0	74·7	78·4	78·0	28·6	76·7	55·8
—	—	68·6	—	68·8	—	—	—	—	—	77·8	45·6	45·8	83·1	91·2	86·4	71·8	38·2	—	62·5
31·6	—	—	—	76·1	—	—	73·1	104·2	3·1	—	50·8	50·8	71·9	—	78·5	—	—	84·3	—
—	—	73·5	72·3	73·8	72·5	101·7	72·0	102·5	1·8	83·2?	43·5	44·6	76·7	93·8	79·7	83·8	29·1	—	50·9?
28·2	59·7	75·9	74·0	76·0	74·1	102·5	73·7	103·0	2·2	70·1	47·5	47·5	77·9	78·0	80·8	78·0	32·4	73·6	44·0?
29·9	59·1	73·8	69·2	73·9	69·3	106·7	69·4	106·6	4·5	79·5	46·3	45·4	80·6	84·0	85·1	72·3	41·2	85·6	52·4
26·9	57·6	72·5	72·5	72·3?	72·4	99·9	72·0?	100·4	0·3?	79·5	—	51·6	73·5	83·6	79·4	84·3	34·5	75·8	35·7
28·0	61·6	—	—	75·5	—	—	68·8?	109·7	6·7?	—	—	—	—	—	—	77·2	26·6	81·2	—
28·0	59·0	—	—	—	—	—	—	96·2?	—	—	—	—	—	—	—	—	—	78·9	—
—	—	—	—	71·7	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
31·0	59·3	74·1	70·9	74·7	71·5	104·5	71·5	104·6	3·3	75·8	42·6	43·6	78·5	82·4	79·7	74·9	39·1	78·3	52·9
31·0	59·9	75·2	73·0	75·1	72·8	103·1	72·9	102·9	2·2	81·6	42·3	42·5	87·1	84·9	91·5	84·7	41·5	84·9	52·2
32·3	58·8	75·9	71·4	76·1	71·6	106·2	71·3	106·7	4·8	72·2	43·7	43·0	85·1	84·9	91·4	86·3	28·2	78·2	61·7
31·1	58·6	71·6	68·5	71·7	68·5	104·6	68·7	104·3	3·0	70·1	53·9	54·5	72·3	77·1	76·2	74·6	36·3	77·2	53·6
29·2	62·1	73·1	69·4	73·2	69·5	105·3	68·6	106·6	4·6	76·0	45·7	46·8	74·0	75·6	80·1	77·7	39·5	79·1	44·2
28·9	62·1	68·3?	73·9	68·6?	74·2	92·4	74·3	92·4?	5·6	—	—	—	81·9	—	87·0	—	—	74·5	—
32·6	65·8	76·8	74·2	77·0	74·4	103·5	73·9	104·2	3·1	75·5	48·6	47·6	79·3	80·0	83·1	83·0	38·3	83·8	53·9
33·9	59·7	73·2	71·6	73·3	71·8	102·2	70·7	103·7	2·6	81·2	47·0	47·7	82·6	80·3	87·8	79·8	38·5	82·5	60·5
28·0	59·2	75·2	70·2	75·4	70·4	107·1	70·8	106·4	4·6?	68·8	51·6	51·6	82·5	83·3	86·8	71·3	28·0	79·5	—
35·1	61·9	77·1	72·1	77·6	72·5	106·9	71·6	108·3	6·0	73·6	42·2	42·6	80·0	78·0	87·8	82·6	35·6	101·4	49·6
28·0	58·8	73·2	71·6	74·3	72·7	102·2	72·7	102·2	1·6	77·6	45·6	46·0	78·7	81·1	83·2	71·2	39·4	79·1	51·2
28·6	61·0	76·0	72·3	76·5	72·8	105·1	71·9	106·4	4·6	71·6	53·4	53·4	81·3	82·8	84·4	—	—	83·9	36·0
27·5	60·2	73·9	76·1	74·0	76·2	97·2	76·1	97·2	− 2·1	73·7	46·8	46·0	82·3	84·9	89·2	75·0	29·5	70·7	63·0
26·8	60·6	75·0	73·9	75·6	74·5	101·5	72·2	104·6	3·4	74·2	49·9	48·8	78·2	74·3	82·6	—	—	85·1	52·7
30·0	56·8	75·6	70·3	75·8	70·5	107·4	70·4	107·7	5·4	80·2	46·1	46·2	77·3	79·7	82·4	76·1	34·8	83·6	51·5
32·1	66·3	76·4	74·4	76·7	74·7	102·6	74·4	103·1	2·3	71·6	46·4	46·4	88·6	87·3	94·4	74·9	28·8	86·8	58·6
29·0	63·4	74·8	73·0	75·1	73·3	102·5	72·9	103·1	2·2	—	45·2	45·1	82·3	—	82·5	67·7	32·4	76·1	41·4
28·1	57·6	71·8	70·1	72·2	70·4	102·5	69·7	103·6	2·5	67·7	47·3	48·4	80·8	83·5	85·1	75·8	32·5	73·9	58·6
—	55·6	71·3	—	71·0	—	—	—	—	—	71·4	57·4	58·0	80·3	80·0	86·0	80·6	24·5	—	44·1
28·9	59·3	72·6	71·0	73·4	71·7	102·3	70·8	103·7	2·6	85·3	42·9	42·8	80·9	84·6	89·3	78·0	37·5	84·8	61·5

TE AND INDICIAL MEASUREMENTS ON 48 CRANIA OF THE FIRST DYNASTY (LENGT

			Angles						
b/fml	*DS/DC*	*SS/SC*	*N∠*	*A∠*	*B∠*	θ_1	θ_2	*P∠*	REMARKS (See Section 2 of '
87·7	56·8	57·9	61°·5	75°·2	43°·3	29°·5	13°·8	89°·0	Teeth worn; P.B. downwards with sharp edge; SR; J equal; ossicle of λ and wo
83·7	50·2	27·6	—	—	—	—	—	—	All teeth except 6 lost; P.B. upwards with sharp, double edge; S equal; J equal
90·9	54·2	37·5	61°·2	75°·0	43°·8	33°·0	10°·8	85°·8	Sagittal suture closing at obelion; 5 teeth lost; P.B. downwards with blunt edge
85·4	54·8?	38·8	67°·0	76°·0	37°·0	25°·6?	11°·4?	87°·4?	Aging; all sutures closing; teeth very worn and 8 lost; P.B. flat with blunt edge
—	50·3	41·0	65°·8	65°·0	49°·2	30°·5	18°·7	83°·7	Aging; coronal and sagittal sutures closed; teeth sound but worn; P.B. flat with
78·4	60·2	37·1	65°·1	73°·8	41°·1	25°·0	16°·1	89°·9	Aging; sutures closing; teeth worn; P.B. downwards with sharp edge; S equal;
78·6	35·7?	25·8	65°·6	68°·9	45°·5	27°·0	18°·5	87°·4	Aging; all sutures closing but teeth perfect; lower 3rd molar L. lost; P.B. flat w
85·8	—	—	57°·6	76°·9	45°·5	34°·3	11°·2	88°·1	Aging; sutures closing; teeth worn but complete; P.B. flat with sharp edge; S e inionic protuberance.
81·8	56·5	55·8	58°·6	74°·5	46°·9	34°·7	12°·2	86°·7	Deep palate; P.B. downwards with rounded edge; S equal; J equal.
73·9	43·9	33·3	58°·8	72°·5	48°·7	33°·3?	15°·4?	87°·9?	Teeth sound but very worn; P.B. flat with sharp edge; SL; JR; flattened obelio
—	—	—	—	—	—	—	—	—	Calvaria and mandible only; tympanic perforation L.; S equal; tripartite interp
79·1	54·8	47·2	60°·6	72°·4	47°·0	30°·5?	16°·5?	88°·9?	Sagittal suture closing at obelion; teeth worn; P.B. downwards with sharp edg epipteric bone R.
85·3	58·7	49·6	60°·1	73°·4	46°·5	31°·6	14°·9	88°·3	Aging; teeth very worn; coronal suture closing; 6 teeth lost from upper jaw; P.
84·1	50·6	27·5	65°·8	70°·0	44°·2	27°·4	16°·8	86°·8	Teeth sound but worn; P.B. downwards with sharp edge; S equal; JR; small ep
98·7	52·8	52·0	65°·0	72°·0	43°·0	33°·1	9°·9	81°·9	No lower 3rd molars; P.B. upwards with sharp edge; S equal; JR.
84·5?	63·4	44·9	64°·3	69°·2	46°·5	32°·1	14°·4	83°·6	Teeth complete but worn; P.B. downwards with rounded edge; SR; JR; complex spine 10 mm. long below and to the left of the inion.
86·2	60·0	36·5	66°·2	71°·1	42°·7	30°·1	12°·6	83°·7	Adult but no upper 3rd molars; teeth worn; tympanic perforation L.; P.B. upwa near asterions.
70·2	42·8	42·9	65°·0	74°·9	40°·1	30°·0	10°·1	85°·0	Teeth worn and 4 lost; no lower 3rd molars; slight tympanic perforation R.; P.
76·7	55·8	61·1	64°·6	74°·8	40°·6	27°·5	13°·1	87°·9	Upper 3rd molars lost and no 3rd molars in lower jaw; P.B. flat with blunt edg
—	62·5	31·9	—	—	—	—	—	81°·5	Aging; sagittal suture closing; teeth worn but complete; P.B. upwards with sha
84·3	—	—	—	—	—	—	—	83°·9	Teeth worn; P.B. flat with blunt edge; J equal.
—	50·9?	41·5?	63°·7	65°·4	50°·9	31°·4?	19°·5?	84°·9?	Aging; sagittal suture closed; teeth worn; P.B. upwards with blunt edge; S equ jecting on left. Only ramus R. of mandible.
73·6	44·0?	29·9?	62°·6	75°·0	42°·4	28°·4?	14°·0?	89°·0?	Aging; sutures closing; teeth worn but complete; slight tympanic perforation L.;
85·6	52·4	40·8	69°·1	67°·3	43°·6	33°·0	10°·6	77°·9	Teeth worn; no 3rd molars in lower jaw; P.B. flat with blunt and double edge;
75·8	35·7	21·0	59°·2	74°·5	46°·3	31°·8	14°·5	89°·0	Outer palate bridge L. and inner palate bridge R.; P.B. downwards with blunt
81·2	—	—	—	—	—	—	—	—	Damaged calvaria and detached facial bones; teeth worn.
78·9	—	—	—	—	—	—	—	—	Aging; calvaria only; sutures closing and vault falling in at obelion; SL; JR; w
—	—	—	—	—	—	—	—	—	Aged; facial bones and ex-occipital missing; teeth of mandible sound out worn; S e
78·3	52·9	62·2	64°·4	71°·9	43°·7	31°·6	12°·1	84°·0	P.B. upwards with sharp edge; SR; J equal; no mandible.
84·9	52·2	53·5	61°·8	70°·2	48°·0	36°·1	11°·9	82°·1	Teeth perfect; P.B. flat with blunt edge; S equal; JL.
78·2	61·7	45·6	62°·9	69°·2	47°·9	35°·1	12°·8	82°·0	Aging; sagittal suture closed; teeth very worn and 2 lost; P.B. upwards with sl with frontal.
77·2	53·6	34·7	68°·8	70°·5	40°·7	26°·2	14°·5	85°·0	Teeth worn; P.B. upwards with sharp edge; S equal; JR; wormians lambdoid R
79·1	44·2	40·5	66°·0	73°·0	41°·0	25°·0	16°·0	89°·0	Teeth perfect; P.B. upwards with sharp edge; SR; JR.
74·5	—	55·8	59°·0	80°·0	41°·0	28°·3?	12°·7?	92°·7?	Aging; coronal suture closing; no 3rd molars in either jaw; left facial and tempo
83·8	53·9	42·0	63°·4	73°·8	42°·8	31°·5	11°·3	85°·1	Teeth worn; slight tympanic perforation R.; P.B. upwards with blunt edge; S 5 mm. high) on left supra-mastoid crest.
82·5	60·5	54·5	64°·2	69°·3	46°·5	30°·9	15°·6	84°·9	Tympanic perforation R. and L.; P.B. upwards with sharp edge; S equal; JR.
79·5	—	—	64°·0	77°·4	38°·6	24°·7	13°·9	91°·3	Sagittal suture closing; teeth worn but all sound; P.B. flat with sharp edge; SR;
101·4	49·6	19·9	68°·6	68°·6	42°·8	29°·5	13°·3	81°·9	Sagittal suture closing at obelion; teeth very worn and many lost; P.B. upwar lambdoid R. and L.; post-coronal depression and very depressed roots of nas
79·1	51·2	32·0	65°·0	69°·2	45°·8	29°·4	16°·4	85°·6	Teeth worn; P.B. upwards with blunt edge; S equal; JL; flattened obelion.
83·9	36·0	31·0	66°·8	70°·9	42°·3	25°·8	16°·5	87°·4	Several teeth lost; sutures beginning to close; P.B. flat with blunt edge; S equal frontal R. and L.; wormians lambdoid R. and L.
70·7	63·0	37·0	65°·1	76°·0	38°·9	35°·7	3°·2	79°·2	Teeth worn and 1 lost; P.B. flat with rounded edge; S equal; JR.
85·1	52·7	43·4	62°·7	75°·5	41°·8	29°·3	12°·5	88°·0	Sagittal suture closing; teeth very worn; P.B. upwards with sharp edge; S equal
83·6	51·5	45·8	63°·2	70°·9	45°·9	30°·4	15°·5	86°·4	P.B. flat with rounded edge; S equal; J equal; epipteric bone L.; single ossicle in s
86·8	58·6	36·3	63°·5	75°·0	41°·5	33°·6	7°·9	82°·9	Teeth worn and no lower 3rd molar R.; P.B. upwards with sharp edge; S equal;
76·1	41·4	38·8	60°·8	72°·8	46°·4	30°·1	16°·3	89°·1	P.B. upwards with sharp edge; S equal; large epipteric bone R.; wormians lamb
73·9	58·6	42·4	66°·8	74°·1	39°·1	28°·2	10°·9	85°·0	Aging; sagittal suture nearly closed; teeth worn; small upper 3rd molar R.; v sharp edge; S equal; JR.
—	44·1	25·8	—	—	—	—	—	81°·5	Young adult; 3rd molars not fully erupted but teeth worn; P.B. flat with blunt
84·8	61·5	73·2	62°·2	72°·3	45°·5	29°·1	16°·4	88°·7	Teeth worn; S equal; JL; P.B. flat with sharp edge; small wormians lambdoid occipital bone.

IAL MEASUREMENTS ON 48 CRANIA OF THE FIRST DYNASTY (Lengths in Mil

	Angles					
N∠	A∠	B∠	θ_1	θ_2	P∠	Remarks (See Section 2 of Text)
1°·5	75°·2	43°·3	29°·5	13°·8	89°·0	Teeth worn; P.B. downwards with sharp edge; SR; J equal; ossicle of λ and wormians lambdo
—	—	—	—	—	—	All teeth except 6 lost; P.B. upwards with sharp, double edge; S equal; J equal.
1°·2	75°·0	43°·8	33°·0	10°·8	85°·8	Sagittal suture closing at obelion; 5 teeth lost; P.B. downwards with blunt edge; SR; JR; pro
7°·0	76°·0	37°·0	25°·6?	11°·4?	87°·4?	Aging; all sutures closing; teeth very worn and 8 lost; P.B. flat with blunt edge; S equal; JL;
5°·8	65°·0	49°·2	30°·5	18°·7	83°·7	Aging; coronal and sagittal sutures closed; teeth sound but worn; P.B. flat with sharp edge; S
5°·1	73°·8	41°·1	25°·0	16°·1	89°·9	Aging; sutures closing; teeth worn; P.B. downwards with sharp edge; S equal; JL; epipteric l
5°·6	68°·9	45°·5	27°·0	18°·5	87°·4	Aging; all sutures closing but teeth perfect; lower 3rd molar L. lost; P.B. flat with sharp edge
5°·6	76°·9	45°·5	34°·3	11°·2	88°·1	Aging; sutures closing; teeth worn but complete; P.B. flat with sharp edge; S equal; JR; asy inionic protuberance.
5°·6	74°·5	46°·9	34°·7	12°·2	86°·7	Deep palate; P.B. downwards with rounded edge; S equal; J equal.
5°·8	72°·5	48°·7	33°·3?	15°·4?	87°·9?	Teeth sound but very worn; P.B. flat with sharp edge; SL; JR; flattened obelion.
—	—	—	—	—	—	Calvaria and mandible only; tympanic perforation L.; S equal; tripartite interparietal with os
5°·6	72°·4	47°·0	30°·5?	16°·5?	88°·9?	Sagittal suture closing at obelion; teeth worn; P.B. downwards with sharp edge; SR; JR; si epipteric bone R.
5°·1	73°·4	46°·5	31°·6	14°·9	88°·3	Aging; teeth very worn; coronal suture closing; 6 teeth lost from upper jaw; P.B. flat with sh
5°·8	70°·0	44°·2	27°·4	16°·8	86°·8	Teeth sound but worn; P.B. downwards with sharp edge; S equal; JR; small epipteric bone L
5°·0	72°·0	43°·0	33°·1	9°·9	81°·9	No lower 3rd molars; P.B. upwards with sharp edge; S equal; JR.
5°·3	69°·2	46°·5	32°·1	14°·4	83°·6	Teeth complete but worn; P.B. downwards with rounded edge; SR; JR; complex lambdoid sut spine 10 mm. long below and to the left of the inion.
5°·2	71°·1	42°·7	30°·1	12°·6	83°·7	Adult but no upper 3rd molars; teeth worn; tympanic perforation L.; P.B. upwards with blunt near asterions.
5°·0	74°·9	40°·1	30°·0	10°·1	85°·0	Teeth worn and 4 lost; no lower 3rd molars; slight tympanic perforation R.; P.B. downwards
5°·6	74°·8	40°·6	27°·5	13°·1	87°·9	Upper 3rd molars lost and no 3rd molars in lower jaw; P.B. flat with blunt edge; S equal; JF
—	—	—	—	—	81°·5	Aging; sagittal suture closing; teeth worn but complete; P.B. upwards with sharp edge; S equ
—	—	—	—	—	83°·9	Teeth worn; P.B. flat with blunt edge; J equal.
5°·7	65°·4	50°·9	31°·4?	19°·5?	84°·9?	Aging; sagittal suture closed; teeth worn; P.B. upwards with blunt edge; S equal; JL; post-c jecting on left. Only ramus R. of mandible.
5°·6	75°·0	42°·4	28°·4?	14°·0?	89°·0?	Aging; sutures closing; teeth worn but complete; slight tympanic perforation L.; P.B. flat with
5°·1	67°·3	43°·6	33°·0	10°·6	77°·9	Teeth worn; no 3rd molars in lower jaw; P.B. flat with blunt and double edge; SR; JR; post
5°·2	74°·5	46°·3	31°·8	14°·5	89°·0	Outer palate bridge L. and inner palate bridge R.; P.B. downwards with blunt edge; SL; sin
—	—	—	—	—	—	Damaged calvaria and detached facial bones; teeth worn.
—	—	—	—	—	—	Aging; calvaria only: sutures closing and vault falling in at obelion; SL; JR; wormians lamb
—	—	—	—	—	—	Aged; facial bones and ex-occipital missing; teeth of mandible sound but worn; S equal; sutures
5°·4	71°·9	43°·7	31°·6	12°·1	84°·0	P.B. upwards with sharp edge; SR; J equal; no mandible.
5°·8	70°·2	48°·0	36°·1	11°·9	82°·1	Teeth perfect; P.B. flat with blunt edge; S equal; JL.
5°·9	69°·2	47°·9	35°·1	12°·8	82°·0	Aging; sagittal suture closed; teeth very worn and 2 lost; P.B. upwards with sharp edge; SI with frontal.
5°·8	70°·5	40°·7	26°·2	14°·5	85°·0	Teeth worn; P.B. upwards with sharp edge; S equal; JR; wormians lambdoid R. and L.
5°·0	73°·0	41°·0	25°·0	16°·0	89°·0	Teeth perfect; P.B. upwards with sharp edge; SR; JR.
5°·0	80°·0	41°·0	28°·3?	12°·7?	92°·7?	Aging; coronal suture closing; no 3rd molars in either jaw; left facial and temporal bones mi
5°·4	73°·8	42°·8	31°·5	11°·3	85°·1	Teeth worn; slight tympanic perforation R.; P.B. upwards with blunt edge; S equal; JL; b 5 mm. high) on left supra-mastoid crest.
5°·2	69°·3	46°·5	30°·9	15°·6	84°·9	Tympanic perforation R. and L.; P.B. upwards with sharp edge; S equal; JR.
5°·0	77°·4	38°·6	24°·7	13°·9	91°·3	Sagittal suture closing; teeth worn but all sound; P.B. flat with sharp edge; SR; JL.
5°·6	68°·6	42°·8	29°·5	13°·3	81°·9	Sagittal suture closing at obelion; teeth very worn and many lost; P.B. upwards with very lambdoid R. and L.; post-coronal depression and very depressed roots of nasal bones.
2°·0	69°·2	45°·8	29°·4	16°·4	85°·6	Teeth worn; P.B. upwards with blunt edge; S equal; JL; flattened obelion.
5°·8	70°·9	42°·3	25°·8	16°·5	87°·4	Several teeth lost; sutures beginning to close; P.B. flat with blunt edge; S equal; J equal; la frontal R. and L.; wormians lambdoid R. and L.
5°·1	76°·0	38°·9	35°·7	3°·2	79°·2	Teeth worn and 1 lost; P.B. flat with rounded edge; S equal; JR.
5°·7	75°·5	41°·8	29°·3	12°·5	88°·0	Sagittal suture closing; teeth very worn; P.B. upwards with sharp edge; S equal; J equal; w
5°·2	70°·9	45°·9	30°·4	15°·5	86°·4	P.B. flat with rounded edge; S equal; J equal; epipteric bone L.; single ossicle in sagittal suture
5°·5	75°·0	41°·5	33°·6	7°·9	82°·9	Teeth worn and no lower 3rd molar R.; P.B. upwards with sharp edge; S equal; JR; epipter
5°·8	72°·8	46°·4	30°·1	16°·3	89°·1	P.B. upwards with sharp edge; S equal; large epipteric bone R.; wormians lambdoid R. and
5°·8	74°·1	39°·1	28°·2	10°·9	85°·0	Aging; sagittal suture nearly closed; teeth worn; small upper 3rd molar R.; very slight ty sharp edge; S equal; JR.
—	—	—	—	—	81°·5	Young adult; 3rd molars not fully erupted but teeth worn; P.B. flat with blunt edge; SR; s
2°·2	72°·3	45°·5	29°·1	16°·4	88°·7	Teeth worn; S equal; JL; P.B. flat with sharp edge; small wormians lambdoid R. and L.; occipital bone.

GTHS IN MILLIMETRES).

f Text)

wormians lambdoid L. No mandible.
ual.
dge; SR; JR; projecting spine below temporal ridge L.
dge; S equal; JL; very projecting angles of mandible.
ith sharp edge; SR; JR.
l; JL; epipteric bone L.
t with sharp edge.
S equal; JR; asymmetrical occiput; massive mastoids and

elion.
erparietal with ossa triangularia only separate.
edge; SR; JR; single ossicles in coronal suture R. and L.;

P.B. flat with sharp edge; SR; JL; ossicle of λ.
l epipteric bone L.

plex lambdoid suture with small wormians; blunt projecting

wards with blunt edge; SR; JL; ossicles in lambdoid suture

P.B. downwards with sharp edge; S equal; JR.
edge; S equal; JR.
sharp edge; S equal; 2 ossicles lambdoid R.

equal; JL; post-coronal depression; facial asymmetry, pro-

L.; P.B. flat with sharp edge; S equal; J equal; no mandible.
ge; SR; JR; post-coronal depression.
nt edge; SL; single wormian lambdoid L.

; wormians lambdoid R. and L. No mandible.
S equal; sutures almost completely closed and vault falling in.

th sharp edge; SR; JR; process of temporal R. articulating

id R. and L.

mporal bones missing; P.B. flat with sharp edge; SR; JR.
e; S equal; JL; blunt projecting spine (10 mm. × 10 mm. by

R.
; SR; JL.

wards with very rounded edge; S equal; J equal; wormians
f nasal bones.

equal; J equal; large processes of temporals articulating with

equal; J equal; wormians lambdoid R. and L.
e in sagittal suture above λ and wormians lambdoid R. and L.
qual; JR; epipteric bone R.
lambdoid R. and L.; left side of mandible only.
R.; very slight tympanic perforation L.; P.B. upwards with

blunt edge; SR; simple interparietal.
doid R. and L.; traces R. and L. of horizontal suture across

NOTE.—*The Definition of the Alveolar Point* (see p. 20 above). By KARL PEARSON.

There has been much vagueness about the localisation of the so-called alveolar point. It culminates probably in Flower's definition*: "*Alveolar Point*, the centre of the anterior margin of the upper alveolar arch." This might appear capable of interpretation had not Flower defined what we term the profile length and then described the *Basialveolar Length* as "Basion to alveolar point, or the most distant part of the anterior margin of the alveolar arch."

Now the most distant part of the anterior margin is rarely in the mesial plane of the skull, but on the anterior face of one or other of the front incisors on either side of the mesial plane. This difficulty has been recognised by certain authorities, and they define the alveolar point to be the point in which a tangent to the border of the alveolar arch on these incisors meets the mesial plane. This is a fairly definite point, but as it usually "hangs in the air" it is not capable of accurate determination in the case of the average skull; it is not at all a suitable definition for the *Knotenpunkt* of a whole series of cranial measurements. We next turn to the classical French craniometricians. Paul Broca in his *Instructions craniologiques et craniométriques*, p. 42, Paris 1875†, writes as follows:

La *région incisive*, comprise entre le nez et la bouche, est formée par la partie de l'arcade alvéolaire qui supporte les dents incisives. On y remarque, sur la ligne médiane, la *suture incisive*, dont l'extrémité supérieure aboutit à l'épine nasale et dont l'extrémité inférieure, correspondant au milieu de l'arcade alvéolaire, se nomme le *point alvéolaire*.

The alveolar point is here defined to be the "lower extremity" of the *incisive suture*, and this is said to correspond to the middle of the "arcade alvéolaire." But where does the incisive suture terminate? Some might hold that it runs as far as the incisive foramen, but no one has ventured to place the alveolar point there. At any rate if it be an *extremity* of the suture we are seeking, it is more likely to be on the posterior than the anterior face of the alveolar process.

Topinard defines the *Point alvéolaire* as the "point le plus antérieur et le plus déclive du bord alvéolaire supérieur‡."

Unfortunately in the great bulk of crania the two characteristics by which Topinard defines the alveolar point are directly contradictory, the most anterior point is not the lowest point.

Turning to English writers we have already referred to Flower. Cunningham§ identifies the Alveolar Point with the Prosthion and defines it as "the centre of the anterior margin of the upper alveolar margin." This leaves us in a state of perplexity as to how a margin can have a margin, much less two margins, as presumably if the upper alveolar margin has an anterior margin it has also a posterior margin. Should the definition be translated "the point where the anterior margin of the upper alveolar process cuts the mesial plane," we must point out that the alveolar margin between any pair of adjacent teeth is extraordinarily vague

* *Catalogue of the R.C. of S. Museum, Homo Sapiens*, Part I, pp. xiii and xviii, 1879.

† See also *Mémoires de la Société d'Anthropologie*, T. II. 2e série, pp. 96 *et seq.*

‡ *Éléments d'Anthropologie Générale*, p. 251, 1885.

§ *Text-book of Anatomy*, p. xvi, Appendix D, 1909.

and often lacks determination, and, according to the personal equation of the observer, may lie anywhere from the geodesic linking up the mid anterior points of the alveolar margin on the incisors to the lowest point of the central process between the incisors. Quain's *Anatomy* is more concise; it reads*:

"Below the nasial aperture is the *intermaxillary suture*; the lower point of this, on the alveolar margin between the central incisors, is named the *alveolar point* (*prosthion*)."

This definition is more or less a return to Broca's where we have reference to the incisive suture and its lower point (if not its "extremity"); the lower point is said to be on the alveolar margin, but then this margin does not terminate the suture and the position of the margin between the central† incisors is the matter under discussion.

Turning to the Germans we naturally consult first the *Frankfurter Verständigung* (*Archiv für Anthropologie*, Bd. XV. 1884), but the only clue to the alveolar point is to be found in the definition of the upper face height (*Obergesichtshöhe*). This we take from nasion to alveolar point. The *Verständigung* defines it as "von der Mitte der Sutura naso-frontalis, bis zur Mitte des Alveolarrandes des Oberkiefers zwischen den mittleren Schneidezähnen" (S. 3). This certainly seems identical with the lowest point of the central process between the middle incisors. The diagrams attached to the *Verständigung* throw no more light on the *accurate* localisation of the alveolar point than those in Broca, Topinard, Cunningham or Quain.

E. Schmidt in his *Anthropologische Methoden*, 1885, gives S. 210 the following definition: "Alveolarpunkt (Point alvéolaire), Kreuzungspunkt des Zahnrandes des Oberkiefers mit der Medianebene." Here again we ask what is the "Zahnrand," the tooth-border, at the mesial plane? R. Martin in his *Lehrbuch der Anthropologie* (S. 514, 1914) sees something of the difficulty and separates the Prosthion from the point to which the upper face height is to be taken. Thus: "Prosthion [pr] = derjenige Punkt am Alveolarrand des Oberkiefers, der in der Mediansagittal-Ebene zwischen den mittleren Schneidezähnen am meisten nach vorn vorragt. Das Prosthion liegt also nicht an dem unteren Ende des zwischen die Schneidezähne vorgeschobenen Knochenteils, sondern an der am meisten nach vorn vorstehenden Stelle seiner Vorderwand. Nur für die Messung der Obergesichtshöhe ist der Messpunkt an die Spitze des genannten Fortsatzes zu legen."

Thus Martin agrees with our interpretation of the *Frankfurter Verständigung*. But he proceeds at once to contradict himself for he defines (S. 553) the *Obergesichtshöhe* (or, the naso-alveolar height for he cites the English term) as the "Geradlinige Entfernung des Nasion vom Prosthion," thus introducing confusion into the *Verständigung*, as well as into his own book. But is it always the case that the most forward projecting point of the alveolar border lies between the middle incisors? We think not. Again, is that border generally distinguishable when it meets the mesial plane? Martin appears to identify it with the most projecting point of the anterior wall of the central process between the middle incisors; but

* *Elements of Anatomy*, Vol. IV. Part I, p. 113, 1915.

† We have heard a stringent critic assert that there cannot be *two* central or *two* mid incisors, which is perfectly correct, but the term is convenient, and what is good enough for Quain is permissible for us.

it is the exact boundary between the alveolar arch and the central process between the mid incisors which is so difficult of determination.

But this is not the greatest difficulty. If there be two points used, the prosthion and the alveolar point—defined, as the Biometricians have defined it since 1895, as the lowest point of the central process between the middle incisor teeth—then the computation of the angles and sides of the fundamental triangle of the skull becomes impossible. Martin would measure the upper face height (according to the *Frankfurter Verständigung*) to the alveolar point, but the profile length to the prosthion*, and he would take his profile angle from the nasion-prosthion line†. It is clear that there could be no orientation of the fundamental triangle (*Gesichtsdreieck*) unless all the measurements were taken to the same point, and accordingly somewhere between his pages 514 and 562, Martin has given up the alveolar point as the extremity of his upper face height. The *Verständigung* itself is very obscure on the measurement of the Upper Face Height and of the Profile Length. We have already given its definition of the former. The latter length is thus defined (S. 4): "von dem am meisten vorspringenden Punkt der Mitte des äusseren Alveolarrandes‡ des Oberkiefers bis zum vorderen Rand des Foramen magnum (in der Medianebene) gemessen." It is by no means clear that the alveolar extremities of the upper face height and the profile length are to be taken to the same point; it looks as if the former was to be taken to the alveolar point and the latter to a point which may or may not be the prosthion.

It was experience of the indefiniteness and positive confusion in this matter and the need for selecting a point from which *all* measurements may be taken, and from which the orientation of the fundamental triangle (*Gesichtsdreieck*) with regard to the Frankfurt horizontal plane may be determined, which led to our selection of the mid-point of the central process between the middle incisors as the alveolar point. As a matter of fact we have preferred to define it not as the mid-point, but as the *lowest* point of the central process when the skull is orientated to the Frankfurt plane§. As far as the fundamental triangle is concerned there is no more valid reason for selecting the anterior than the posterior face of the alveolar arch. I regret only that we did not take palate lengths from this alveolar point rather than from a point on the posterior alveolar margin as vague as the prosthion on the anterior alveolar margin.

* *Loc. cit.* S. 550. † *Loc. cit.* S. 560.

‡ There seems here to be some confusion of the alveolar border and alveolar arch; for if *Alveolarrand* signifies the alveolar border or lower margin of the alveolar arch, this is presumably a curve which can have only one middle point, *i.e.* intersects the mesial plane in a single point. What is therefore the sense of introducing the words "the most projecting point"? We can't choose a point within a point. It would appear therefore either that border is used for arch, and the point in question may lie anywhere on the incisive suture where we reach the greatest distance from the basion, or that the *Alveolarrand* is an area in the neighbourhood of the alveolar border on which area we are to choose the point most remote from the basion. How this area is separated from the anterior face of the alveolar arch we do not know, nor have we found it anywhere stated. Another alternative, namely that the mid-point of the alveolar margin is invariably the point furthest removed from the basion, would not only render the definition tautological, but it involves a dogma so obviously erroneous that it need not be discussed.

§ The importance of this will be readily realised by the reader who has examined the central process in numerous crania; central processes occur which are wedge-shaped V rather than semi-circular ∪.

I realise that objection may be taken to measuring the profile angle as we have done, probably on the ground that the nasion-prosthion line is nearer to the profile slope in the living than is the nasion-alveolar point line. But both are so far off the living profile—owing to the flesh thickness of the lip—that the small horizontal projected distance from prosthion to alveolar point makes little difference compared with the great divergence between the cranial and living profiles. This argument, indeed, might justify the use of Cuvier's facial line to the tip of the incisors* as better in its turn than using the nasion-prosthion line.

If further justification be needed for my choice in 1895 of the alveolar point as on the central process of the middle incisors I think it may be found in the *Entente Internationale* of 1906, resulting from the Monaco Congress†, for the unification of craniometric measures. Internationally this concordat should replace the *Frankfurter Verständigung.* Therein we read under the definition of the *Diamètre naso-alvéolaire,* our upper face height, that the lower extremity of this measurement is "le point le plus inférieur du bord‡ alvéolaire entre les deux incisives médianes et supérieures."

Again for the *Diamètre alvéolo-basilaire,* our profile length, the anterior extremity is the "point alvéolaire, ou point médian du bord antérieur de l'arcade alvéolaire." Although the Frankfurt Concordat had been on trial for more than twenty years the distinguished craniometricians of the Committee of the Monaco Congress do not appear to have agreed on a common alveolar point for all their measurements or, if they did, they describe it in one of their definitions as the lowest point, and in the other as simply the mid-point on the alveolar margin, which itself has a very dubious existence in the mesial plane!

It is very clear that one and all these definitions contain fossilised fragments of earlier conceptions; they can hardly be the product of craniometricians, who have had to struggle with the actual difficulty of determining the prosthion or alveolar point on long series of crania, nor of those who realise the importance of eliminating personal equation.

We might define the prosthion as the point in which the geodesic§ on the anterior face of the alveolar arch between the mid-points of the anterior faces of the middle incisors at the anterior alveolar border meets the incisive suture; and again the alveolar point as the point in which the incisive suture or the mesial plane meets the lowest element (relative to the Frankfurt horizontal) of the central process between the middle incisors. In other words it is the lowest point of the section of the above central process by the mesial plane. If the section be straight edged ___/, then we must select its middle point as alveolar point.

We believe that our alveolar point admits of far readier determination than the prosthion, and when it cannot be ascertained owing to absorption of the alveolar borders, then also no prosthion is available.

* Or better still a tangent from the nasion to the alveolar arch in the mesial plane.

† *L'Anthropologie,* Tom. XVII. pp. 563—7.

‡ Does "bord" here signify "arcade"? For if not there is again tautology if the "bord" is a curve meeting the mesial plane.

§ In practice the geodesic is determined by drawing with a fine pencil the shortest line (curved not straight) between the mid-points of the anterior alveolar border on the middle incisors.

ON THE MOMENTS OF THE HYPERGEOMETRICAL SERIES.

By Professor V. ROMANOVSKY.

(University of Taskend.)

In a paper of the same title Prof. K. Pearson (*Biometrika*, Vol. XVI. pp. 157—162) has given a finite difference relation, which enables us to write down the successive moments of a hypergeometrical series about its mean. This relation is very elegant and leaves nothing to desire if we evaluate the moments one after another. But it does not enable us to write down any desired moment without knowing the preceding ones. The aim of this article is to provide a formula, which permits us to obtain the moments of a hypergeometrical series independently.

Let $\alpha = -r$, $\beta = -qn$, $\gamma = pn - r + 1$, and

$$F(\alpha, \beta, \gamma, x) = A\left[1 + \frac{r}{1}\frac{qn}{pn-r+1}x + \frac{r(r-1)}{1.2}\frac{qn(qn-1)}{(pn-r+1)(pn-r+2)}x^2 + \ldots\right],$$

where

$$A = \frac{pn(pn-1)\ldots(pn-r+1)}{n(n-1)\ldots(n-r+1)}.$$

Here the coefficient of x^t is the probability of drawing t individuals without some mark in a sample of size r drawn from a population of n containing pn marked and qn non-marked individuals.

Let $E(t)$ stand for the mathematical expectation of t*. Then

$$E(t) = rq,$$

and it is easy to see by successive differentiations of the function F, that

$$E[t(t-1)\ldots(t-s+1)] = \left[\frac{d^sF}{dx^s}\right]_{x=1} \quad \ldots\ldots\ldots\ldots\ldots\ldots(1),$$

or

$$U_s = \left[\frac{d^sF}{dx^s}\right]_{x=1} \quad \ldots\ldots\ldots\ldots\ldots\ldots(1\ bis),$$

if we put

$$E[t(t-1)\ldots(t-s+1)] = U_s \quad \ldots\ldots\ldots\ldots\ldots\ldots(2).$$

The function $F(\alpha, \beta, \gamma, x)$ satisfies the well-known differential equation

$$x(1-x)\frac{d^2F}{dx^2} + (\gamma - \delta x)\frac{dF}{dx} - \alpha\beta F = 0,$$

where

$$\delta = \alpha + \beta + 1 = 1 - r - qn.$$

* If $F(-r, -qn, pn-r+1, x)$ be written $a_0 + a_1x + a_2x^2 + \ldots + a_tx^t + \ldots + a_rx^r$, we have $S_{t=0}^{t=r}(a_t) = 1$ and $E(f(t)) = S_{t=0}^{t=r}(a_t f(t))$, or $E(f(t))$ is the mean value of $f(t)$ ($t=0$ to r) weighted with the probability of drawing t unmarked individuals in the sample of r.

Differentiating this equation s times we have

$$x(1-x)\frac{d^{s+2}F}{dx^{s+2}}+s(1-2x)\frac{d^{s+1}F}{dx^{s+1}}-s(s-1)\frac{d^sF}{dx^s}+(\gamma-\delta x)\frac{d^{s+1}F}{dx^{s+1}}$$
$$-s\delta\frac{d^sF}{dx^s}-\alpha\beta\frac{d^sF}{dx^s}=0.$$

Put here $x=1$ and taking regard to (1) and (2), we have

$$(\gamma-\delta-s)\,U_{s+1}=[\alpha\beta+s\delta+s(s-1)]\,U_s,$$

or
$$(n-s)\,U_{s+1}=(r-s)(qn-s)\,U_s \quad \ldots\ldots\ldots\ldots(3).$$

Now $U_1=rq$ and thus, with the aid of the above relation, we find

$$U_s=\frac{r(r-1)\ldots(r-s+1)\,qn(qn-1)\ldots(qn-s+1)}{n(n-1)\ldots(n-s+1)} \quad \ldots\ldots\ldots\ldots(4).$$

The sth moment μ_s of our hypergeometrical series is defined by the relation

$$\mu_s=E(t-rq)^s.$$

This may be written

$$\mu_s=E[K_s(t)]$$
$$=E\left[K_s(0)+t\frac{\Delta K_s(0)}{1}+t(t-1)\frac{\Delta^2K_s(0)}{2!}+\ldots\right.$$
$$\left.+t(t-1)\ldots(t-s+1)\frac{\Delta^sK_s(0)}{s!}\right],$$

if we put

$$(t-rq)^s=K_s(t),$$

and apply the interpolation formula of Newton.

Therefrom

$$\mu_s=K_s(0)+\frac{\Delta K_s(0)}{1}U_1+\frac{\Delta^2K_s(0)}{2!}U_2+\ldots+\frac{\Delta^sK_s(0)}{s!}U_s \quad \ldots\ldots\ldots(5).$$

Here

$$\Delta^hK_s(0)=[\Delta^hK_s(t)]_{t=0}=\Delta^h0^s-C_s^1rq\Delta^h0^{s-1}+C_s^2(rq)^2\Delta^h0^{s-2}-\ldots \quad \ldots(6),$$

as it is easy to see.

Let us put

$$\frac{\Delta^h0^\lambda}{h!}=h_\lambda \quad \ldots\ldots\ldots\ldots(7).$$

Applying the relation

$$\Delta^h0^\lambda=h^\lambda-C_h^1(h-1)^\lambda+C_h^2(h-2)^\lambda-\ldots+(-1)^{h-1}C_h^1\,.\,1^\lambda \quad \ldots\ldots\ldots(8),$$

we easily deduce that

$$h_\lambda=h\,.\,h_{\lambda-1}+(h-1)_{\lambda-1} \quad \ldots\ldots\ldots\ldots(9),$$

for all positive integer values of h and λ. It must be remarked that

$$h_h=1 \text{ and } h_\lambda=0 \text{ for } h>\lambda \quad \ldots\ldots\ldots\ldots(10).$$

With the aid of these relations we easily find the values of h_λ for different values of h and λ. We place here a short table of these values.

Table of values of h_λ for h, $\lambda = 1, 2, \ldots, 10$.

h	λ									
	1	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*
1	1	1	1	1	1	1	1	1	1	1
2	—	1	3	7	15	31	63	127	255	511
3	—	—	1	6	25	90	301	966	3025	9330
4	—	—	—	1	10	65	350	1701	7770	34105
5	—	—	—	—	1	15	140	1050	6951	42525
6	—	—	—	—	—	1	21	266	2646	22827
7	—	—	—	—	—	—	1	28	462	5880
8	—	—	—	—	—	—	—	1	36	750
9	—	—	—	—	—	—	—	—	1	45
10	—	—	—	—	—	—	—	—	—	1

Now put

$$\frac{\Delta^h K_s(0)}{h!} = L_{h,s} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(11).$$

Then, according to (6) and (7),

$$L_{h,s} = h_s - C_s^1 h_{s-1} rq + C_s^2 h_{s-2} (rq)^2 - \ldots + (-1)^s C_s^s h_0 (rq)^s,$$

or, in consequence of (10),

$$L_{h,s} = h_s - C_s^1 h_{s-1} rq + C_s^2 h_{s-2} (rq)^2 - \ldots + (-1)^{s-h} C_s^{s-h} (rq)^{s-h} \quad \ldots(12).$$

The polynomials $L_{h,s}$ for $s \leqq 10$ can be written down by the aid of the table. For greater values of s we must recur to (8) or (9). Putting, for example, $s = 5$, we shall have

$$\begin{aligned}
L_{0,5} &= -(rq)^5, \\
L_{1,5} &= 1 - 5rq + 10(rq)^2 - 10(rq)^3 + 5(rq)^4, \\
L_{2,5} &= 15 - 35rq + 30(rq)^2 - 10(rq)^3, \\
L_{3,5} &= 25 - 30rq + 10(rq)^2, \\
L_{4,5} &= 10 - 5rq, \\
L_{5,5} &= 1.
\end{aligned}$$

We see from (5) and (11) that

$$\mu_s = L_{0,s} + L_{1,s} U_1 + L_{2,s} U_2 + \ldots + L_{s,s} U_s \quad \ldots\ldots\ldots\ldots\ldots(13),$$

which is the relation I proposed to establish.

For example, taking $s = 5$, we find

$$\begin{aligned}
\mu_5 = &-(rq)^5 + [1 - 5rq + 10(rq)^2 - 10(rq)^3 + 5(rq)^4] \frac{rqn}{n} \\
&+ [15 - 35rq + 30(rq)^2 - 10(rq)^3] \frac{r(r-1)\,qn(qn-1)}{n(n-1)} \\
&+ [25 - 30rq + 10(rq)^2] \frac{r(r-1)(r-2)\,qn(qn-1)(qn-2)}{n(n-1)(n-2)} \\
&+ [10 - 5rq] \frac{r(r-1)(r-2)(r-3)\,qn(qn-1)(qn-2)(qn-3)}{n(n-1)(n-2)(n-3)} \\
&+ \frac{r(r-1)(r-2)(r-3)(r-4)\,qn(qn-1)(qn-2)(qn-3)(qn-4)}{n(n-1)(n-2)(n-3)(n-4)}.
\end{aligned}$$

Putting in (13) $n = \infty$, we shall find the sth moment for a binomial $(p+q)^r$ about its mean

$$\mu_s = L_{0,s} + L_{1,s}\, rq + L_{2,s}\, r(r-1)\, q^2 + \ldots + L_{s,s}\, r(r-1)\ldots(r-s+1)\, q^s \ldots (14),$$

where the coefficients $L_{h,s}$ keep their values. This relation may be transformed into a more condensed form which is indicated in my note on the moments of a binomial $(p+q)^r$ (see *Biometrika*, Vol. XV. formula (5)).

It is not difficult to establish for hypergeometrical moments the relations

$$\lim (\mu_{2s}/\mu_2^{\,s}) = \frac{1}{\sqrt{\pi}} \int_{-\infty}^{+\infty} x^{2s} e^{-x^2} dx,$$

$$\lim (\mu_{2s+1}/\mu_2^{\,s+\frac{1}{2}}) = \frac{1}{\sqrt{\pi}} \int_{-\infty}^{+\infty} x^{2s+1} e^{-x^2} dx,$$

for $n \to \infty$ and $r \to \infty$. Therefore, according to the theorem of Tchebycheff-Markoff, we conclude that, for a hypergeometrical series, the probability of inequalities

$$t_1 \leqq \frac{t - rq}{\sqrt{\mu_2}} \leqq t_2,$$

for any given t_1 and t_2, tends to $\frac{1}{\sqrt{\pi}} \int_{t_1}^{t_2} e^{-x^2} dx$ for $n \to \infty$ and $r \to \infty$, i.e. for n and r sufficiently great the hypergeometrical distribution is near to the normal curve, as was long ago indicated by Laplace. (*Mémoires de mathématique...présentés à l'Académie*, Tom. VI. p. 6, 1774.)

[*Note.* The two methods, i.e. deducing the higher moments in succession by Pearson's formula and deducing the higher moment directly by Romanovsky's formula, were applied to obtain the fifth moment when

$$n = 100,\ p = \tfrac{1}{10},\ q = \tfrac{9}{10} \text{ and } r = 10.$$

The first gave

$$\mu_2 = \tfrac{9}{10},\ \mu_3 = -\tfrac{288}{539},\ \mu_4 = 2{\cdot}156{,}743 \text{ and } \mu_5 = -4{\cdot}122{,}411 \text{ in succession.}$$

The second gave directly

$$\mu_5 = -4{\cdot}122{,}410.$$

There was very little difference in the time or labour required to reach μ_5, using an arithmometer. But the former gave with no greater expenditure of labour, μ_2, μ_3 and μ_4, which also would generally be required, and the work would be very much longer by the second method, if all the moments were required. If, however, Romanovsky's method were used to check the μ_5 of the first method it would at the same time check all the earlier moments, and this indicates a decided field of usefulness for the formula. ED.]

PROBABILITY INTEGRALS FOR A HYPERGEOMETRIC SERIES.

By Professor BURTON H. CAMP, Ph.D.

(1) If a bag contains pn black balls and qn other balls, where $p+q=1$, the probabilities of obtaining $r, r-1, \ldots, 1, 0$ black balls in one drawing of r balls at a time are given by the terms of the finite hypergeometric* series

$$\frac{y^{(r)}}{n^{(r)}}\left\{1+\frac{r^{(1)}}{1!}\frac{n-y}{y-r+1}+\frac{r^{(2)}}{2!}\frac{(n-y)(n-y-1)}{(y-r+1)(y-r+2)}+\ldots\right\} \quad \ldots\ldots\ldots\text{(i)},$$

where $y=pn$, $n-y=qn$, $y^{(r)}=y(y-1)\ldots(y-r+1)$, etc. The probability of obtaining at least $s=r-t$ black balls is given by the sum of the first $(t+1)$ terms of (i), which may be written in the form

$$S_{t+1}(y)=\frac{1}{n^{(r)}}\left\{y^{(r)}+\frac{r^{(1)}}{1!}y^{(r-1)}(n-y)^{(1)}+\ldots+\frac{r^{(t)}}{t!}y^{(s)}(n-y)^{(t)}\right\} \quad \ldots\ldots\text{(ii)}$$

$$=u(0)+u(1)+\ldots+u(t).$$

It is important to find a ready method of calculating the sum in (ii). This sum is what is meant by the probability integral for the hypergeometric. It is easily seen that, in limiting ourselves to the consideration of the sum of the first $(t+1)$ terms, instead of including also the sum of the last $(t+1)$ terms, we are really not losing generality, for the first sum would become the second if the values of p and q were interchanged, and the sum of all the terms is unity.

(2) The general plan and notation of this paper will follow fairly closely those of an earlier one, and we shall make use of certain of the results given there†. In particular; if A denote the tail of any frequency function $f(\xi)$ such that $f \geqq 0$ if $a \leqq \xi \leqq b$, and $f=0$ elsewhere, and if ‡

$$\Sigma_a^b f(\xi)\,\Delta\xi=1, \quad A=\Sigma_a^\tau f(\xi)\,\Delta\xi=\{f(a)+f(a+\Delta\xi)+\ldots+f(\tau)\}\,\Delta\xi \quad \text{(iii)},$$

then we have, using the symbol "$\cong$" to mean "is approximately equal to," and Sheppard's§ notation,

$$A\cong\phi(0)\left\{\frac{\Delta v}{2}+E\Delta v+\frac{\sigma}{2}\left(\frac{1-\alpha_x}{z_x}\right)\right\},$$

* Cf. W. P. Elderton, *Frequency Curves and Correlation*, Layton, London, 1906, p. 37.

† B. H. Camp, "Probability Integrals for the Point Binomial," *Biometrika*, Vol. xvi. pp. 163—171.

‡ The reader should note that, in the notation to be used, Σ_a^τ is a sum which includes the values at the limits a and τ.

§ *Tables for Statisticians and Biometricians*, 2nd Edition, Part i. 1924, pp. 2—11.

where, in the order suitable for computation,

$$v=\tau-\xi,\ \phi(v)=f(\tau-v),\ \Delta v=\Delta\xi,\ Q=\frac{\phi(\Delta v)}{\phi(0)},\ R=\sqrt{Q\,\frac{\phi(\Delta v)}{\phi(2\Delta v)}},$$

$$c_1=-\log_e QR,\ c_2=\log_e R,\ \frac{1}{\sigma}=\sqrt{2c_2}\,\Delta v,\ x=c_1\sigma,$$

$$E\cong\frac{c_1}{12}\left(1+\frac{3-x^2}{60\sigma^2}+\frac{15-10x^2+x^4}{2520\sigma^4}\right) \quad \text{.....(iv).}$$

Sometimes certain terms of (iv) may be neglected, and then

$$A\cong\frac{\phi(0)\,\Delta v}{1-Q^{\Delta v}} \quad \text{.....(v).}$$

(3) *Special Case* (*Q small*). It is only when Q is small that (v) suffices, but when the stump is far from the mode this often occurs. We may then apply (v) directly to (ii) and obtain

$$S_{t+1}(y)\cong\frac{u(t)}{1-Q(t)} \quad \text{.....(vi),}$$

where $u(t)$ is the first term in the stump, viz.,

$$u(t)=\frac{r^{(t)}}{n^{(r)}\,t!}\,y^{(s)}\,(n-y)^{(t)},\ \text{and}\ Q(t)=\frac{t\,(y-s)}{(s+1)\,(n-y-t+1)}.$$

Sometimes it is worth while computing a few terms of the tail before applying (vi). When three extra terms are computed, $S_{t+1}(y)$ takes the form

$$S_{t+1}(y)=u(t-3)\left\{\frac{1}{1-Q(t-3)}+\frac{1}{Q(t-2)}+\frac{1}{Q(t-2)}\cdot\frac{1}{Q(t-1)}+\frac{1}{Q(t-2)}\cdot\frac{1}{Q(t-1)}\cdot\frac{1}{Q(t)}\right\} \quad \text{.....(vi \textit{bis});}$$

where

$$Q(t-3)=\frac{(t-3)\,(y-s-3)}{(s+4)\,(n-y-t+4)},\qquad \frac{1}{Q(t-2)}=\frac{(s+3)\,(n-y-t+3)}{(t-2)\,(y-s-2)},\ \text{etc.}$$

(4) *General Case.* In order to be able to use for the general case methods analogous to those employed for the point binomial, it is necessary to show that an analogous transformation can be made of the sum S_{t+1}. The result will not be, as in the case of the point binomial, an integral, but instead another sum. However, this new sum will be found to possess a certain advantage due to the fact that the point binomial integral is its limiting form.

First, let us difference (ii) with respect to y, taking $\Delta y=1$, and noting the formulae:

$$\Delta(n-y)^{(t)}=-t\,(n-y-1)^{(t-1)},$$

$$\Delta y^{(s)}\,(n-y)^{(t)}=sy^{(s-1)}\,(n-y)^{(t)}-t\,(y+1)^{(s)}\,(n-y-1)^{(t-1)};$$

$$\Delta S_{t+1}(y)=\frac{1}{n^{(r)}}\left\{ry^{(r-1)}+\frac{r^{(2)}}{1!}\,y^{(r-2)}\,(n-y)^{(1)}+\frac{r^{(3)}}{2!}\,y^{(r-3)}\,(n-y)^{(2)}+\dots\right.$$

$$+\frac{r^{(t+1)}}{t!}\,y^{(s-1)}\,(n-y)^{(t)}-\frac{r^{(1)}}{0!}\,(y+1)^{(r-1)}\,(n-y-1)^{(0)}-\frac{r^{(2)}}{1!}\,(y+1)^{(r-2)}\,(n-y-1)^{(1)}-\dots$$

$$\left.-\frac{r^{(t)}}{(t-1)!}\,(y+1)^{(s)}\,(n-y-1)^{(t-1)}\right\} \quad \text{.....(vii).}$$

Now write (vii) successively for $y=\alpha$, $y=\alpha+1$, ..., $y=\alpha+\beta$, and add. The result is

$$\Delta S_{t+1}(\alpha)+\Delta S_{t+1}(\alpha+1)+\ldots+\Delta S_{t+1}(\alpha+\beta)$$
$$=\frac{1}{n^{(r)}}\frac{r^{(t+1)}}{t!}\Sigma_{\xi=\alpha}^{\xi=\alpha+\beta}\xi^{(s-1)}(n-\xi)^{(t)}(\Delta\xi=1)$$
$$-\frac{1}{n^{(r)}}\left[\frac{r^{(t)}}{(t-1)!}\xi^{(s)}(n-\xi)^{(t-1)}+\frac{r^{(t-1)}}{(t-2)!}\xi^{(s+1)}(n-\xi)^{(t-2)}+\ldots+\frac{r^{(1)}}{0!}\xi^{(r-1)}(n-\xi)^{(0)}\right]_{\xi=\alpha}^{\xi=\alpha+\beta+1}$$
.....(viii).

In this let us now put $\alpha=s-1$, $\alpha+\beta=y-1$. Then the left hand side of (viii) equals $S_{t+1}(y)$, for by (ii) $S_{t+1}(s-1)=0$. On the right the expression in brackets vanishes at the lower limit $\xi=s-1$, for each term involves $(s-1)$ with an index greater than $s-1$, omitting the trivial case where $t=0$. Thus we have

$$S_{t+1}(y)=\frac{1}{n^{(r)}}\frac{r^{(t+1)}}{t!}\Sigma_{\xi=s-1}^{\xi=y-1}\xi^{(s-1)}(n-\xi)^{(t)}\Delta\xi+\theta(y),$$

where

$$\theta(y)=-\frac{1}{n^{(r)}}\left\{\frac{r^{(t)}}{(t-1)!}y^{(s)}(n-y)^{(t-1)}+\frac{r^{(t-1)}}{(t-2)!}y^{(s+1)}(n-y)^{(t-2)}+\ldots+\frac{r^{(1)}}{0!}y^{(r-1)}(n-y)^{(0)}\right\}$$
.....(ix).

In order to sum the series $\theta(y)$, we may repeat the method just used. Difference with respect to y, getting $\Delta\theta(y)$.

Then write $\Delta\theta(y)$ for $y=\alpha$, $y=\alpha+1$, ... $y=\alpha+\beta$, and add, obtaining

$$\Delta\theta(\alpha)+\Delta\theta(\alpha+1)+\ldots+\Delta\theta(\alpha+\beta).$$

If we then substitute $\alpha=s-1$, $\alpha+\beta=y-1$, we obtain $\theta(y)$, for $\theta(s-1)=0$, as already noted. The details, though somewhat lengthy, offer no difficulty, and are omitted. We obtain, finally,

$$\theta(y)=-\frac{1}{n^{(r)}}\frac{r^{(t+1)}}{(t-1)!}\Sigma_{\xi=s-1}^{\xi=y-1}\xi^{(s-1)}(n-\xi)^{(t-1)}(\Delta\xi=1)+\theta_1(y),$$

where

$$\theta_1(y)=\frac{1}{n^{(r)}}\Sigma_{i=1}^{i=t-1}\frac{r^{(t+1-i)}}{(t-1-i)!}y^{(s-1+i)}(n-y)^{(t-1-i)}(\Delta i=1)$$ .....(x).

Now $\theta_1(y)$ is exactly like $\theta(y)$, except for sign and certain changes in the indices. The result of continually repeating this process may therefore be inferred, and the final result is

$$S_{t+1}(y)=\frac{r^{(t+1)}}{n^{(r)}}\Sigma_{\xi=s-1}^{\xi=y-1}\xi^{(s-1)}\left(\frac{(n-\xi)^{(t)}}{t!}-\frac{(n-\xi)^{(t-1)}}{(t-1)!}+\ldots\pm 1\right)$$ .....(xi).

The series in (xi) is in the form

$${}_mC_t-{}_mC_{t-1}+\ldots\pm{}_mC_0$$ .....(xii),

where ${}_mC_t$ is the number of combinations of m things taken t at a time, etc. In order to sum it, we write Pascal's theorem for t, $t-1$, etc., and add, as follows:

$$\begin{aligned}
{}_mC_t &= {}_{m-1}C_t + {}_{m-1}C_{t-1},\\
-{}_mC_{t-1} &= -{}_{m-1}C_{t-1} - {}_{m-1}C_{t-2},\\
{}_mC_{t-2} &= {}_{m-1}C_{t-2} + {}_{m-1}C_{t-3},\\
&\vdots \qquad\quad \vdots \qquad\quad \vdots\\
\text{Sum} &= {}_{m-1}C_t
\end{aligned}$$ .....(xiii).

Insert this result in (xi):

$$S_{t+1}(y) = \frac{r^{(t+1)}}{n^{(r)}\,t!} \sum_{\xi=s-1}^{\xi=y-1} \xi^{(s-1)}\,(n-\xi-1)^{(t)}\,(\Delta\xi = 1) \quad \ldots\ldots\ldots\text{(xiv)}.$$

This is the formula for the hypergeometric which is analogous to (xiv) in my paper on the Point Binomial. It is of the form (iii), and, if we now apply to it the approximation (iv), which, as in my previous paper, is expected to be good only when s and t are large, we have as

METHOD I (*Indirect Approximation*): $S_{t+1}(y) = u(t)\frac{s}{y}\rho$, where $u(t)$ is as in (vi), and, in the order suitable for computation,

$$Q(t) = \frac{(y-s)(n-y+1)}{(y-1)(n-y-t+1)}, \quad R = \sqrt{Q\frac{(y-2)(n-y-t+2)}{(y-s-1)(n-y+2)}},$$

$$c_1 = -\log_e QR, \quad c_2 = \log_e R, \quad \frac{1}{\sigma} = \sqrt{2c_2}, \quad x = c_1\sigma,$$

$$E \text{ as in (iv)}, \quad \rho = \tfrac{1}{2} + \frac{\sigma}{2}\left(\frac{1-\alpha_x}{z_x}\right) + E \quad \ldots\ldots\ldots\text{(xv)}.$$

As in the case of the point binomial, it is necessary that $c_1 \geqq 0$, and this can always be secured by a proper statement of the problem initially. That is, if c_1 should ever come out negative, one would know that the tail of the hypergeometric as chosen had produced in (xiv) a function of ξ whose mode was within the limits of the summation indicated in that formula. The original problem should therefore be restated so that the opposite end of the hypergeometric is selected as the tail, and then c_1 will be found to be positive. The mode of the function of ξ that occurs is easily found to be at the point $\xi = n(s-1)/(r-1) - 1$.

Example 1. Find the sum of the first 46 terms of the hypergeometric for which $n = 360$, $r = 90$, $p = \frac{1}{3}$.

Here $y = 120$, $t = 45$, $n - y = 240$, $Q = \dfrac{(75)(241)}{(196)(119)} = \cdot 774\ 9528$,

$R = \sqrt{Q\dfrac{(118)(197)}{(74)(242)}}$, $c_1 = 0{\cdot}251\ 9864$, $c_2 = 0{\cdot}002\ 9662\ 102$,

$\dfrac{1}{\sigma} = 0{\cdot}077\ 0222$, $x = 3{\cdot}271\ 607$, $E = 0{\cdot}020\ 983$, $\rho = 4{\cdot}1920$,

$S = 1{\cdot}5720$, $u(t) = 0{\cdot}0001\ 1553$;

and the true value of this sum, found by computing and adding the terms, is 0·0001 1525.

(5) METHOD II (*Direct Approximation*): As in the case of the point binomial one may approximate to the sum of the first $(t+1)$ terms of a hypergeometric by applying the formulae of Section 2 (iv) directly to the series. The formulae resulting from this process will now be given, and then, after a few illustrative examples, the relative merits of the two methods will be discussed:

$$S_{t+1} = u(t), \quad Q = \frac{t(y-s)}{(s+1)(n-y-t+1)}, \quad R = \sqrt{Q\frac{(s+2)(n-y-t+2)}{(t-1)(y-s-1)}},$$

c_1, c_2, σ, x, E, and ρ as in (xv)……(xvi).

Example 2. Find the sum in Example 1 by (xvi). We obtain

$$c_1 = 0{\cdot}951\ 358,\quad c_2 = 0{\cdot}031\ 2456,\quad \frac{1}{\sigma} = 0{\cdot}249\ 982,\quad x = 3{\cdot}80570,\quad E = 0{\cdot}078\ 341,$$

$$\rho = 1{\cdot}56828,\quad S = 0{\cdot}000\ 115\ 253, \text{ and the true value is } 0{\cdot}000\ 115\ 254.$$

(6) The above examples have been worked with all the precision necessary to exhibit the closeness of approximation of which the formulae are capable. This has involved the use of inconvenient tables* to find $\frac{1}{2}(1-\alpha)/z$, and the necessity of finding the constants to several decimal places. For many purposes, however, as great accuracy as that obtained, for example, by (xvi), is not required. What is especially needed is a method which will give results in error by at most one or two per cent., and which will involve a minimum of computation. This last condition is very important, for we are here considering approximate formulae, and one always has the option of discarding the formulae and computing the several terms of the series outright. The approximate formulae, therefore, must not be too cumbersome to employ readily. Both of the methods given above admit of some modification, which will now be noted, and the succeeding examples will illustrate the accuracy which may be obtained with a minimum of effort.

First write, in both cases, $E \cong c_1/12$, and so :

Indirect Method I: $$S_{t+1} = \frac{s}{y}\,u(t)\left\{\tfrac{1}{2} + \frac{c_1}{12} + \frac{\sigma}{2}\,\frac{(1-\alpha_x)}{z_x}\right\} \quad \text{.....(xvii),}$$

Direct Method II: $$S_{t+1} = u(t)\ \left\{\tfrac{1}{2} + \frac{c_1}{12} + \frac{\sigma}{2}\,\frac{(1-\alpha_x)}{z_x}\right\} \quad \text{.....(xviii).}$$

In both cases, find Q and R^2 by a computing machine, and then $\log_e Q$ and $\log_e R^2$ from a five place table of natural† logarithms. Obtain σ to four significant figures by the use of tables of reciprocals and square roots, and find the ratio $\frac{1}{2}(1-\alpha)/z$ from Sheppard's table‡.

Example 3. Do Example 1 as just indicated. By Method I, $S = 0{\cdot}000\ 116$; by Method II, $S = 0{\cdot}000\ 115$; the true value is 0·000 115.

Example 4. Let $n = 900$, $r = 450$, $p = \frac{2}{3}$, $t = 120$. By Method I, $S = 0{\cdot}000\ 0145$; by Method II, $S = 0{\cdot}000\ 0144$; the true value is 0·000 0145.

Example 5. Let $n = 900$, $r = 450$, $p = \frac{2}{3}$, $t = 110$. By Method I, $S = 1{\cdot}00 \times 10^{-8}$; by Method II, $S = 1{\cdot}01 \times 10^{-8}$; the true value is $1{\cdot}02 \times 10^{-8}$.

Example 6. Let $n = 240$, $r = 90$, $p = \frac{1}{3}$, $t = 45$. By Method I, $S = 0{\cdot}0000\ 2285$; by Method II, $S = 0{\cdot}0000\ 2282$; the true value is 0·0000 2282.

Example 7. Let $n = 360$, $r = 90$, $p = \frac{1}{5}$, $t = 72$. By Method I, $S = 0{\cdot}002\ 783$; by Method II, $S = 0{\cdot}002\ 756$; the true value is 0·002 754.

* Burgess, J., *Transactions of the Royal Society of Edinburgh*, 1897, Vol. xxxix. pp. 257—321.

† E.g., B. O. Pierce, *A Short Table of Integrals*, 2nd Ed., 1910, Ginn and Co., Boston, p. 130.

‡ Unfortunately, this table is so arranged that the ratio is given poorly in the range, $4 \leqq x < 4{\cdot}50$. A tabulation of the ratio itself would be very helpful here, as with the point binomial approximation. Such a table, accurate to five significant figures, is now in course of preparation.

Example 8. Let $n = 900$, $r = 90$, $p = \frac{1}{3}$, $t = 72$. By Method I, $S = 0{\cdot}006\ 535$; by Method II, $S = 0{\cdot}006\ 428$; the true value is $0{\cdot}006\ 428$.

Example 9. Let $n = 900$, $r = 90$, $p = \frac{1}{2}$, $t = 45$. By Method I, $S = 0{\cdot}000\ 422$; by Method II, $S = 0{\cdot}000\ 414$; the true value is $0{\cdot}000\ 414$.

A critical survey of these examples will indicate that the direct method (II) gives better results. There is no essential difference between the two methods as regards difficulty of computation. Is there, therefore, justification for the development of the indirect method? This question may be answered after a brief consideration of the point binomial case. There the indirect method was preferred, not because it gave closer results—indeed, in the published examples it gave less close results—but because it gave results which were close enough, and it involved less computation. It was, for large values of r (n in that paper), the more *efficient* approximation, efficiency being understood to depend partly on the closeness of the approximation, and partly on the simplicity of the formulae. Now, if we let n become infinite in our present indirect formulae, they will approach as limits the corresponding indirect formulae for the point binomial; the sum (xiv) of this paper will approach the integral (xiv) of that one. Therefore, in border line cases, that is, where n/r is large, it is to be expected, and it is in fact true that, without serious loss of accuracy, the formulae of this method may be simplified, so as to render them analogous to the point binomial formulae, and so as to make them easier to use in computing. On the other hand, although the direct method formulae are susceptible of an analogous transformation, going over into the direct method formulae for the point binomial, they do not result in a corresponding simplification. The simplified formulae will now be given and illustrated.

(7) *Use of Indirect Method when n/r is large.* As before, use tables of reciprocals and squares and find the following constants accurately to four significant figures:

$$c_1 = \frac{s-1}{y-1} - \frac{t}{n-y-t+1} + \tfrac{1}{2}\left[\left(\frac{s-1}{y-1}\right)^2 + \left(\frac{t}{n-y-t+1}\right)^2\right],$$

$$\frac{1}{\sigma^2} = \frac{-1}{y-1} + \frac{1}{n-y-t+1} - \frac{1}{n-y+2} + \frac{1}{y-s-1}, \quad x = c_1\sigma.$$

Then

$$S_{t+1} = u(t)\,\frac{s}{y}\left(\tfrac{1}{2} + \frac{\sigma}{2}\,\frac{1-\alpha_x}{z_x}\right) \qquad \text{.....(xix).}$$

Example 10. Use the above method in Example 9. We obtain $n/r = 10$, $S = 0{\cdot}000\ 424$; by I, Ex. 9, $S = 0{\cdot}000\ 422$: the true value is $0{\cdot}000\ 414$. Thus there is no essential difference in the approximation given by Method I when these simpler formulae for the constants are used. Since r is only 90, neither result is very close to the true value. As with the point binomial, the success of the method depends on r being large *. This method is especially valuable when r is much larger, as in the following case.

* I have intentionally included in both my papers examples where the methods do not apply very well, in order that the reader may have some idea of the range of their usefulness.

Example 11. Let $n = 4500$, $r = 450$, $p = \frac{2}{3}$, $t = 120$. Here also $n/r = 10$. $S = 0{\cdot}000\ 7965$; the true value is $0{\cdot}000\ 7980$. The proportional error is only about $0{\cdot}19$ of one per cent., one-twelfth as large as that in Example 10. Also, the last place is in some doubt, due to the inadequacy of the tables for the ratio, $\frac{1}{2}(1-\alpha)/z$.

(8) *Summary.* In order that the reader may have the benefit of such experience as the author has had in computing from the above formulae, some suggestions as to when the various formulae should be used will now be given. These suggestions are admittedly both tentative and proximate, for the experience on which they are based is not exhaustive, and because, when the accuracy of a formula depends on four variables (n, r, p, s), it is seldom possible to frame a simple, rigorous rule of procedure.

(*a*) If greater accuracy is desired than can be obtained by the abbreviated methods of Section (6), it is probably not worth while using the formulae at all unless r is at least as large as 300. This would not be true if better auxiliary tables were available*, but in our present state it is probably better to compute the various terms of the tail by successive multiplications, beginning with the term at the stump.

(*b*) If $pq > \frac{1}{10}$, and $r < 400$, use Method II of Section (6). If $pq > \frac{1}{10}$ and $r > 400$, use Method II of Section (6) for values of $n < 10r$, Method I as modified in Section (7) for higher values of n. The proportional errors will probably not exceed one- or two-tenths of one per cent. in the extreme cases, and will usually be much smaller. They grow small as pq and r are increased.

(*c*) Formula (vi *bis*), or even (vi), is always to be used, however, when the Q of Method II, which is the same as the Q of (vi), is small, say less than one-third. A case in which Q is a little greater than one-third ($0{\cdot}3743$) is given in Examples 2 and 3. The error in S as given by the method of Example 2 is $0{\cdot}000\ 000\ 001$; by II of Example 3 it is $0{\cdot}000\ 000\ 02$; by (vi *bis*) it is $0{\cdot}000\ 000\ 19$; and by (vi) it is $0{\cdot}000\ 002\ 26$. The proportional error by (vi *bis*) is $0{\cdot}0017$, and by (vi) is $0{\cdot}02$.

Finally, it is a pleasure to record my indebtedness to Professor Karl Pearson for proposing the problem of this paper and for helpful suggestions made during the course of its preparation.

* Cf. footnote ‡ on p. 65 above.

DETERMINATION OF $\int_0^\theta \cos^{n+1}\theta\, d\theta$ FOR LARGE VALUES OF n, AND ITS APPLICATION TO THE PROBABILITY INTEGRAL OF SYMMETRICAL FREQUENCY CURVES.

By JOHN WISHART, M.A., B.Sc.

The Complete Integral $\int_0^{\pi/2} \cos^{n+1}\theta\, d\theta$ has been tabled from $n=-1$ to 103*. The Incomplete Integral

$$\int_0^\theta \cos^{n+1}\theta\, d\theta = \tfrac{1}{2}\int_0^x (1-x)^{\frac{1}{2}n} x^{-\frac{1}{2}}\, dx \quad \text{on putting} \quad x=\sin^2\theta$$

$$= \tfrac{1}{2} B_x(\tfrac{1}{2}n+1, \tfrac{1}{2}), \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(1)$$

and so can be found from Tables of the Incomplete B-function, which are in progress. But such tables must of necessity be limited, stopping, say, at $p=q=50$. It still remains a problem then to determine the integral under consideration for large values of n. It is the object of this paper to show how our function can be evaluated to a very great degree of accuracy by expanding in Incomplete Normal Moment functions. The method was first suggested by Professor Pearson as a means of evaluating the incomplete Γ- and B-functions†, but it was found after tables of the moment functions had been made that it was not accurate enough when n was small. Illustrations of the method in the case of the Incomplete B-function appear in *Tracts for Computers*, No. VII, but the accuracy there achieved leaves something to be desired.

As will be shown later, the curve of which (1) is the integral is one of the well-known Pearson frequency types, and hence what is generally required is the ratio of the incomplete to the complete integral.

Let
$$\tfrac{1}{2} + I_\theta(n+1) = \frac{\int_{-\pi/2}^{\theta} \cos^{n+1}\theta\, d\theta}{\int_{-\pi/2}^{+\pi/2} \cos^{n+1}\theta\, d\theta}.$$

Then
$$I_\theta(n+1) = \frac{\int_0^\theta \cos^{n+1}\theta\, d\theta}{2\int_0^{\pi/2} \cos^{n+1}\theta\, d\theta} \quad \ldots\ldots\ldots\ldots\ldots\ldots(2)$$

$$= \frac{\int_0^y (1-y^2)^{\frac{1}{2}n}\, dy}{2\int_0^1 (1-y^2)^{\frac{1}{2}n}\, dy} \quad \ldots\ldots\ldots\ldots\ldots\ldots(3)$$

on putting $y=\sin\theta$.

* *Biometrika*, Vol. XI. p. 377

† *Biometrika*, Vol. VI. pp. 66—8.

Now $$\int_0^1 y^{2m-1}(1-y^2)^{p-1}\,dy = \tfrac{1}{2}\cdot\frac{\Gamma(m)\,\Gamma(p)}{\Gamma(m+p)}.$$

In this put $m=\frac{1}{2}$ and $p=\frac{1}{2}n+1$.

$$\therefore\ I_\theta(n+1)=\frac{1}{\sqrt{\pi}}\cdot\frac{n+1}{n}\cdot\frac{\Gamma\{\frac{1}{2}(n+1)\}}{\Gamma(\frac{1}{2}n)}\int_0^y(1-y^2)^{\frac{1}{2}n}\,dy.$$

Again $$\Gamma(\tfrac{1}{2}n)\,\Gamma\{\tfrac{1}{2}(n+1)\}=\frac{\sqrt{\pi}}{2^{n-1}}\Gamma(n)$$

and we shall expand the Γ-functions by Stirling's theorem. Thus:

$$I_\theta(n+1)=\frac{1}{2^{n-1}}\left(1+\frac{1}{n}\right)\frac{\Gamma(n)}{\{\Gamma(\frac{1}{2}n)\}^2}\int_0^y e^{-\frac{1}{2}n\,(y^2+\frac{1}{2}y^4+\frac{1}{3}y^6+\frac{1}{4}y^8+\ldots)}\,dy \quad\ldots\ldots\ldots(4)$$

$$=\left(1+\frac{1}{n}\right)\left(1-\frac{1}{4n}+\frac{1}{32n^2}+\frac{5}{128n^3}\right)\frac{1}{\sqrt{2\pi}}\int_0^x e^{-\frac{x^2}{2}}\cdot e^{-\frac{x^4}{4n}-\frac{x^6}{6n^2}-\frac{x^8}{8n^3}}\,dx,$$

on putting $x=\sqrt{n}\,.\,y$ in the integral. Terms in $\frac{1}{n^4}$ and higher powers are neglected.

$$\therefore\ I_\theta(n+1)=\left(1+\frac{3}{4n}-\frac{7}{32n^2}+\frac{9}{128n^3}\right)\frac{1}{\sqrt{2\pi}}$$

$$\times\int_0^x e^{-\frac{x^2}{2}}\left[1-\frac{x^4}{4n}-\frac{x^6}{6n^2}+x^8\left(\frac{1}{32n^2}-\frac{1}{8n^3}\right)+\frac{x^{10}}{24n^3}-\frac{x^{12}}{384n^3}\right]dx.$$

Now the rth Incomplete Normal Moment function is

$$m_r(x)=\frac{1}{\sqrt{2\pi}}\int_0^x x^r e^{-\frac{1}{2}x^2}\,dx\Big/(r-1)(r-3)\ldots\ldots 2 \text{ or } 1,$$

and so in terms of these

$$I_\theta(n+1)=\left(1+\frac{3}{4n}-\frac{7}{32n^2}+\frac{9}{128n^3}\right)m_0(x)-\left(\frac{3}{4n}+\frac{9}{16n^2}-\frac{21}{128n^3}\right)m_4(x)$$

$$-\left(\frac{5}{2n^2}+\frac{15}{8n^3}\right)m_6(x)+\left(\frac{105}{32n^2}-\frac{1365}{128n^3}\right)m_8(x)+\frac{315}{8n^3}m_{10}(x)-\frac{3465}{128n^3}m_{12}(x)\ \ldots(5)$$

$$=c_0\,.\,m_0(x)-c_4\,.\,m_4(x)-c_6\,.\,m_6(x)+c_8\,.\,m_8(x)+c_{10}\,.\,m_{10}(x)-c_{12}\,.\,m_{12}(x)\ \ldots\ldots(6),$$

where

$$\left.\begin{aligned}
c_0&=1+\frac{\cdot75}{n}-\frac{\cdot21875}{n^2}+\frac{\cdot0703125}{n^3}\\
c_4&=\frac{\cdot75}{n}+\frac{\cdot5625}{n^2}-\frac{\cdot1640625}{n^3}\\
c_6&=\frac{2\cdot5}{n^2}+\frac{1\cdot875}{n^3}\\
c_8&=\frac{3\cdot28125}{n^2}-\frac{10\cdot6640625}{n^3}\\
c_{10}&=\frac{39\cdot375}{n^3}\\
c_{12}&=\frac{27\cdot0703125}{n^3}
\end{aligned}\right\}\ \ldots\ldots\ldots\ldots\ldots(7).$$

In this expression $x = \sqrt{n}\sin\theta$. $m_0(x)$ is got from the table of $\frac{1}{2}(1+\alpha)$* by subtracting ·5, and the functions $m_r(x)$ from $r = 1$ to 10 are tabled to seven figures*. A supplementary table giving $m_{12}(x)$ appears below. It was calculated by integrating twice by parts and combining the results, which gives the relation

$$m_{12}(x) = \tfrac{1}{2}\{m_8(x) + m_{10}(x)\} - \frac{x^9}{1890}\left(1 + \frac{2x^2}{11}\right) z_x,$$

where z_x is the ordinate of the normal curve, given by Sheppard's tables. In a few cases where the last figure was in doubt, the integration was carried further, and the function expressed in terms of $m_0(x)$, known to more places than seven. Second differences only are necessary for interpolation into this table, for although the third differences are considerable, none is greater than ·0006, and this, when combined with the appropriate coefficient, is insignificant to the order of accuracy we desire.

This result (equation (6)) enables the function $I_\theta(n+1)$ to be calculated with great rapidity when once the coefficients c_0 etc. are known, and provided that n is reasonably large. Table I gives the values of these coefficients for values of n from 100 to 400, proceeding by twos, i.e. for values which would come outside the Tables of the Incomplete B-function. For still higher values of n, c_0, c_4, c_6 and c_8 can be calculated from (7), dropping the terms in $\dfrac{1}{n^3}$. It is found that for n above 400, the terms in $\dfrac{1}{n^3}$ are insignificant in the seventh place, and to a first approximation

$$I_\theta(n+1) = \left(1 + \frac{3}{4n} - \frac{7}{32n^2}\right) m_0(x) - \left(\frac{3}{4n} + \frac{9}{16n^2}\right) m_4(x) - \frac{5}{2n^2}\, m_6(x) + \frac{105}{32n^2}\, m_8(x) \ldots (8).$$

Now the equation to the symmetrical frequency curve of limited range (Pearson, Type IIa) is

$$z = z_0\left(1 - \frac{x^2}{a^2}\right)^m \dagger.$$

If in this we put $y = \dfrac{x}{a}$ and $m = \frac{1}{2}n$ it becomes

$$z = z_0(1 - y^2)^{\frac{1}{2}n},$$

and hence the probability integral of this type of curve is

$$\frac{\int_{-1}^{y}(1-y^2)^{\frac{1}{2}n}\, dy}{\int_{-1}^{+1}(1-y^2)^{\frac{1}{2}n}\, dy} = \tfrac{1}{2} + I_\theta(n+1),$$

from (3).

So that the function determined by equation (6) is the probability integral of the Type IIa curve, when added to $\frac{1}{2}$. The diagram illustrates the form the curve

* *Tables for Statisticians*, pp. 2—8 and 22—3.

† The type $z = z_0\left(1 + \dfrac{x^2}{a^2}\right)^{-m}$ reduces to the same integral form if we take $x/a = \tan\theta$; no further probability integral is required.

takes for large values of n, and the probability integral is the ratio of the shaded portion of the area to the whole area. The examples which will now be given in

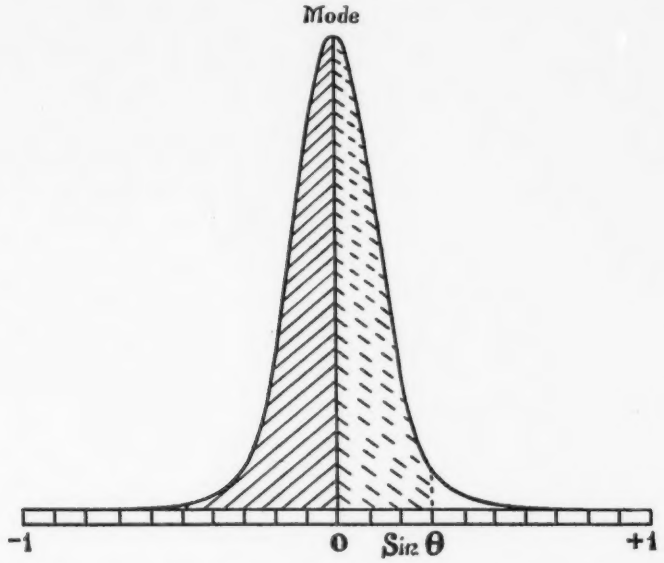


illustration of the theory will show that the approximation (6) determines $I_\theta(n+1)$ to seven figure accuracy up to a distance from the mode (taken as origin) of three times the standard deviation. This is very nearly the whole area. The value of the "tail" is given by $\frac{1}{2} - I_\theta(n+1)$.

Example I. To evaluate $I_\theta(101)$.

$$\sigma = \frac{1}{\sqrt{n+3}} = \cdot 0985.$$

(*a*) Let us integrate from the mode to $\theta = \sin^{-1} \cdot 3$.

Then $x = \sqrt{n} \sin\theta = 3$. We have at once from equation (6) and using Table I

$$\begin{aligned} I_{\cdot 3}(101) = {} & 1\cdot 0074\ 782\ m_0(3) - \cdot 0075\ 561\ m_4(3) \\ & - \cdot 0002\ 519\ m_6(3) + \cdot 0003\ 175\ m_8(3) \\ & + \cdot 0000\ 394\ m_{10}(3) - \cdot 0000\ 271\ m_{12}(3) \\ = {} & \cdot 5023\ 7911 \quad - \cdot 0033\ 6600 \\ & \quad 8933 \qquad\quad 9413 \\ & \quad 745 \qquad\quad 308 \\ & \cdot 5024\ 7589 \quad \cdot 0034\ 6321 \\ & \quad 34\ 6321 \\ & \cdot 4990\ 1268 \end{aligned}$$

$\therefore\ I_{\cdot 3}(101) = \cdot 4990\ 127$ to seven places.

As a check on this result, a value, correct to at least seven figures, was obtained from 16 terms of the "parts" reduction formula. Since

$$\int_0^y (1-y^2)^{\frac{1}{2}n}\, dy = \int_0^y (1+y)^{\frac{1}{2}n}\,(1-y)^{\frac{1}{2}n}\, dy$$

the function we are evaluating is seen to be equal to $\dfrac{B_y(51,51)}{B(51,51)}$, taken from the mode as origin, and when we remember that our range is ± 1 from the mode and the integral of the "parts" reduction formula is measured from the start of the curve, we see that

$$I_{\cdot 3}(101) = \frac{1}{2} - \frac{B_{\cdot 35}(51,51)}{B(51,51)}$$

$$= \frac{1}{2} - \frac{\Gamma(102)}{\Gamma(52)\,\Gamma(51)}(\cdot 35)^{51}(\cdot 65)^{50}\left(1 + \frac{50}{52}\cdot\frac{7}{13}\left\{1 + \frac{49}{53}\cdot\frac{7}{13}\left\{\ldots\ldots\left\{1 + \frac{36}{66}\cdot\frac{7}{13}\right\}\right\}\ldots\right\}\right),$$

the Remainder Integral being insignificant to this order of accuracy,

$$= \frac{1}{2} - \frac{\Gamma(102)}{\Gamma(52)\,\Gamma(51)}(\cdot 35)^{51}(\cdot 65)^{50} \times 1{\cdot}9984\ 3246.$$

This expression was evaluated using Tables of the Logarithms of the Gamma-Function* and Briggs' Logarithms to 12 places.

We have

	159·974 325 0285
	$\overline{24}$·747 470 2618
	$\overline{10}$·645 667 8322
	0·300 689 4751
	127·668 152 5976
	130·673 719 9211
	$\overline{4}$·994 432 6765

$$\therefore\ I_{\cdot 3}(101) = \tfrac{1}{2} - \cdot 0009\ 8726 = \cdot 4990\ 127 \text{ to seven places,}$$

which shows that our value by moment functions is correct to seven figures. A further attempt was made towards evaluating the integral, using Weddle's Rule (24 ordinates). This gave a value ·4990 128.

The actual value of $\int_0^\theta \cos^{101}\theta\, d\theta$, where $\theta = \sin^{-1}\cdot 3$, is, of course, obtained by multiplying the value found for $I_\theta(101)$ by $2\int_0^{\pi/2}\cos^{101}\theta\, d\theta$, which is tabled.

(*b*) Let us now integrate over a smaller range, say to $\theta = \sin^{-1}\cdot 2$. Then $x = \sqrt{n} \times \cdot 2 = 2$. The coefficients are the same as before, but we now require to look out the moment coefficients of $x = 2$. We find numerically:

+ ·4808 1884	− ·0017 0233
1406	2774
59	12
·4808 3349	·0017 3019
17 3019	
·4791 0330	

$$\therefore\ I_{\cdot 2}(101) = \cdot 4791\ 033 \text{ to seven places,}$$

* *Tracts for Computers*, No. VIII.

which is again found to be identical to this number of places with the result deduced by parts.

Example II. Consider now the case of a higher power.

Let it be proposed to find $I_\theta(401)$. Now the point of inflection of the curve is given by $\sin\theta = y = \dfrac{1}{\sqrt{n-1}}$, so for the case considered in Example I the points of inflection are $\pm\cdot1$ from the mode, and in this example $\pm\cdot05$. Sometimes an approximate formula is found to be unreliable in the neighbourhood of a point of inflection. To examine this difficulty, let us take

(a) Integration to $\theta = \sin^{-1}\cdot05$, i.e. to a distance from the mode a little greater than the standard deviation, $\sigma = \cdot0498$.

Then $$x = \sqrt{n}\sin\theta = 1.$$

Equation (8) may be used, the extra terms of (5) being found to be of no significance in the seventh place. We then have

$$\begin{aligned} I_{\cdot05}(401) &= 1\cdot0018\ 736\ m_0(1) - \cdot0018\ 785\ m_4(1) \\ &\quad - \cdot0000\ 156\ m_6(1) + \cdot0000\ 205\ m_8(1) \\ &= \cdot3419\ 8430 \qquad - \cdot0000\ 3516 \\ &\qquad\quad 3520 \qquad\qquad\quad 4 \\ &\qquad \overline{\cdot3419\ 491} \quad \text{to seven places,} \end{aligned}$$

the result to this number of places being, in fact, got from the first two terms only.

Two checks were here employed. Weddle's Rule, applied to 24 ordinates, gave the result ·3419 491. Then the integral was expanded and integrated term by term, as follows:

From (4) $$I_\theta(n+1) = \frac{1}{2^{n-1}}\left(1+\frac{1}{n}\right)\frac{\Gamma(n)}{\{\Gamma(\frac{1}{2}n)\}^2}\int_0^y (1-y^2)^{\frac{1}{2}n}\,dy$$

$$= \frac{2\cdot005}{2^{400}}\,\frac{\Gamma(400)}{\{\Gamma(200)\}^2}\left[(\cdot05) - \frac{200}{3}(\cdot05)^3 + \frac{200\,.\,199}{2!\,5}(\cdot05)^5 - \ldots\right].$$

We need seven terms of the series for seven figure accuracy.

$$\therefore\ I_{\cdot05}(401) = \frac{\Gamma(400)}{2^{400}\{\Gamma(200)\}^2} \times \cdot0857\ 675\ 241.$$

We have

866·204 354 1864	745·191 717 2888
$\bar{2}$·933 322 8727	120·411 998 2656
865·137 677 0591	
865·603 715 5544	
$\bar{1}$·533 961 5047	

$$\therefore\ I_{\cdot05}(401) = \cdot3419\ 491 \text{ to seven places,}$$

so that all three methods yield, to this number of places, identical results.

(*b*) To integrate now over the range 3σ.

Let $\theta = \sin^{-1}\cdot144$. $\therefore\ x = 20 \times \cdot144 = 2\cdot88$.

The values of the moment functions of 2·88 were obtained by interpolation, using Everett's Central Difference Formula

$$z_\theta = \phi z_0 + \theta z_1 - \tfrac{1}{6}\theta\phi\{(\phi+1)\,\delta^2 z_0 + (\theta+1)\,\delta^2 z_1\}, \qquad \theta = \cdot8,\ \phi = \cdot2.$$

Consideration of the fourth differences throughout the table of moment functions shows us that, though in some cases they are considerable, the next term in Everett's Formula, taken in conjunction with the maximum coefficient of Table I, is negligible to our order of accuracy.

We then have

$$\begin{aligned} I_{\cdot144}(401) = {} & 1\cdot0018\ 736 \times \cdot4980\ 11624 \\ & - \cdot0018\ 785 \times \cdot4296\ 294 \\ & - \cdot0000\ 156 \times \cdot3463\ 230 \\ & + \cdot0000\ 205 \times \cdot2476\ 137 \\ = {} & \cdot4981\ 373 \text{ to seven places.} \end{aligned}$$

The value found by Weddle's Rule was also ·4981 373, while integration term by term, a much more lengthy process this time, since 20 terms are required for seven figure accuracy, gave

$$\begin{aligned} I_{\cdot144}(401) &= \frac{\Gamma(400)}{2^{400}\{\Gamma(200)\}^2} \times \cdot1249\ 425\ 775 \\ &= \cdot4981\ 373. \end{aligned}$$

Conclusion. The foregoing examples prove fairly conclusively that results, correct to seven places, can be obtained by expanding the function (2) in terms of the tabled seven figure moment functions. The first example would normally be obtained from a Table of the Incomplete *B*-function, when available, but it was selected as a case of extremely low indices. It is clear that the larger n is, the easier does the computation become, for then equation (5) reduces to the comparatively simple form of (8).

TABLE I.

Values of the Coefficients.

n	c_0	c_4	c_6	c_8	c_{10}	c_{12}
100	1·0074782	·0075561	·0002519	·0003175 −	·0000394	·0000271
102	1·0073320	74069	2421	3053	371	255 +
4	1·0071914	72634	2328	2939	350 +	241
6	1·0070561	71254	2241	2831	331	227
8	1·0069257	69925 +	2158	2728	313	215 −
110	1·0068002	68645 +	2080	2632	296	203
112	1·0066790	67412	2006	2540	280	193
4	1·0065622	66221	1936	2453	266	183
6	1·0064493	65072	1870	2370	252	173
8	1·0063403	63962	1807	2292	240	165 −
120	1·0062348	62890	1747	2217	228	157
122	1·0061329	61852	1690	2146	217	149
4	1·0060342	60849	1636	2078	207	142
6	1·0059386	59877	1584	2013	197	135 +
8	1·0058461	58936	1535 −	1952	188	129
130	1·0057563	58024	1488	1893	179	123
132	1·0056693	57140	1443	1837	171	118
4	1·0055849	56283	1400	1783	164	113
6	1·0055029	55451	1359	1732	157	108
8	1·0054233	54643	1320	1682	150 −	103
140	1·0053460	53858	1282	1635 +	143	99
142	1·0052709	53095 +	1246	1590	138	95 −
4	1·0051978	52354	1212	1547	132	91
6	1·0051267	51633	1179	1505 +	127	87
8	1·0050576	50932	1147	1465 +	121	84
150	1·0049903	50250 −	1117	1427	117	80
152	1·0049248	49585 +	1087	1390	112	77
4	1·0048609	48938	1059	1354	108	74
6	1·0047987	48308	1032	1320	104	71
8	1·0047381	47693	1006	1287	100	69
160	1·0046790	47094	981	1256	96	66
162	1·0046213	46510	957	1225 +	93	64
4	1·0045651	45940	934	1196	89	61
6	1·0045101	45384	911	1167	86	59
8	1·0044566	44842	890	1140	83	57
170	1·0044042	44312	869	1114	80	55 +
172	1·0043531	43794	849	1088	77	53
4	1·0043031	43289	829	1064	75 −	51
6	1·0042543	42795 −	811	1040	72	50 −
8	1·0042066	42312	792	1017	70	48
180	1·0041599	41840	775 −	994	68	46
182	1·0041143	41378	758	973	65 +	45 −
4	1·0040696	40927	741	952	63	43
6	1·0040259	40485 −	726	932	61	42
8	1·0039832	40053	710	912	59	41
190	1·0039413	39629	695 +	893	57	39
192	1·0039003	39215 −	681	875 +	56	38
4	1·0038602	38809	667	857	54	37
6	1·0038208	38412	653	840	52	36
8	1·0037823	38022	640	823	51	35 −
200	1·0037445 +	37640	627	807	49	34

TABLE I (*continued*).

Values of the Coefficients.

n	c_0	c_4	c_6	c_8	c_{10}	c_{12}
202	1·0037075+	·0037266	·0000615−	·0000791	·0000048	·0000033
4	1·0036712	36900	603	776	46	32
6	1·0036356	36540	591	761	45+	31
8	1·0036007	36188	580	747	44	30
210	1·0035665−	35842	569	733	43	29
212	1·0035329	35502	558	719	41	28
4	1·0034999	35169	548	706	40	28
6	1·0034675+	34843	538	693	39	27
8	1·0034358	34522	528	680	38	26
220	1·0034046	34207	518	668	37	25+
222	1·0033739	33898	509	656	36	25−
4	1·0033439	33594	500−	644	35+	24
6	1·0033143	33296	491	633	34	23
8	1·0032853	33003	482	622	33	23
230	1·0032567	32715−	474	612	32	22
232	1·0032287	32432	466	601	32	22
4	1·0032011	32154	458	591	31	21
6	1·0031740	31881	450+	581	30	21
8	1·0031474	31612	443	571	29	20
240	1·0031212	31348	435+	562	28	20
242	1·0030954	31088	428	553	28	19
4	1·0030701	30832	421	544	27	19
6	1·0030452	30581	414	535+	26	18
8	1·0030206	30333	408	527	26	18
250	1·0029965+	30090	401	518	25+	17
252	1·0029727	29850+	395−	510	25−	17
4	1·0029494	29615−	389	502	24	17
6	1·0029264	29383	383	494	23	16
8	1·0029037	29154	377	487	23	16
260	1·0028814	28929	371	479	22	15+
262	1·0028594	28708	365+	472	22	15+
4	1·0028378	28490	360	465−	21	15−
6	1·0028165−	28275−	354	458	21	14
8	1·0027955−	28063	349	451	20	14
270	1·0027748	27855−	344	445−	20	14
272	1·0027544	27649	339	438	20	13
4	1·0027343	27447	334	432	19	13
6	1·0027145+	27248	329	426	19	13
8	1·0026950+	27051	324	420	18	13
280	1·0026758	26857	320	414	18	12
282	1·0026568	26666	315+	408	18	12
4	1·0026381	26478	311	402	17	12
6	1·0026197	26292	306	397	17	12
8	1·0026015+	26109	302	391	16	11
290	1·0025836	25929	298	386	16	11
292	1·0025659	25751	294	381	16	11
4	1·0025485−	25575+	290	375+	15+	11
6	1·0025313	25402	286	370	15+	10
8	1·0025143	25231	282	365+	15−	10
300	1·0024976	25062	278	361	15−	10

TABLE I (*continued*).

Values of the Coefficients.

n	c_0	c_4	c_6	c_8	c_{10}	c_{12}
302	1·0024810	·0024896	·0000275 −	·0000356	·0000014	·0000010
4	1·0024647	24732	271	351	14	10
6	1·0024486	24570	268	347	14	9
8	1·0024328	24410	264	342	13	9
310	1·0024171	24252	261	338	13	9
312	1·0024016	24096	257	334	13	9
4	1·0023863	23942	254	329	13	9
6	1·0023712	23790	251	325	12	9
8	1·0023563	23640	248	321	12	8
320	1·0023416	23492	245 −	317	12	8
322	1·0023271	23346	242	313	12	8
4	1·0023127	23202	239	309	12	8
6	1·0022986	23059	236	306	11	8
8	1·0022846	22918	233	302	11	8
330	1·0022707	22779	230	298	11	8
332	1·0022571	22641	227	295 −	11	7
4	1·0022435 +	22505 +	225 −	291	11	7
6	1·0022302	22371	222	288	10	7
8	1·0022170	22239	219	284	10	7
340	1·0022040	22107	217	281	10	7
342	1·0021911	21978	214	278	10	7
4	1·0021784	21850 −	212	275 −	10	7
6	1·0021658	21723	209	272	10	7
8	1·0021534	21598	207	268	9	6
350	1·0021411	21474	205 −	265 +	9	6
352	1·0021289	21352	202	262	9	6
4	1·0021169	21231	200	259	9	6
6	1·0021050 +	21112	198	257	9	6
8	1·0020933	20994	195 +	254	9	6
360	1·0020816	20877	193	251	8	6
362	1·0020702	20761	191	248	8	6
4	1·0020588	20647	189	245 +	8	6
6	1·0020475 +	20534	187	243	8	6
8	1·0020364	20422	185 −	240	8	5 +
370	1·0020254	20311	183	238	8	5 +
372	1·0020145 +	20202	181	235 +	8	5 +
4	1·0020038	20094	179	233	8	5 +
6	1·0019931	19987	177	230	7	5 +
8	1·0019826	19881	175 +	228	7	5 +
380	1·0019722	19776	173	225 +	7	5 −
382	1·0019619	19672	172	223	7	5 −
4	1·0019516	19569	170	221	7	5 −
6	1·0019415 +	19468	168	218	7	5 −
8	1·0019315 +	19367	166	216	7	5 −
390	1·0019216	19268	165 −	214	7	5 −
392	1·0019118	19169	163	212	7	4
4	1·0019021	19072	161	210	6	4
6	1·0018925 +	18975 +	160	208	6	4
8	1·0018830	18880	158	205 +	6	4
400	1·0018736	18785 +	157	203	6	4

TABLE II.

Values of the 12th Incomplete Normal Moment Function.

x	$m_{12}(x)$	δ^2	x	$m_{12}(x)$	δ^2
0·0	·0000000		2·5	·0317107	+23061
0·1	·0000000		2·6	·0429539	25057
0·2	·0000000		2·7	·0567028	26157
0·3	·0000000		2·8	·0730674	26170
0·4	·0000000		2·9	·0920490	24977
0·5	·0000000		3·0	·1135283	22562
0·6	·0000000		3·1	·1372638	19005
0·7	·0000000		3·2	·1628998	14482
0·8	·0000001	+ 4	3·3	·1899840	9265
0·9	·0000006	7	3·4	·2179947	3657
1·0	·0000018	30	3·5	·2463711	− 2004
1·1	·0000060	68	3·6	·2745471	7391
1·2	·0000170	152	3·7	·3019840	12221
1·3	·0000432	315	3·8	·3281988	16270
1·4	·0001009	601	3·9	·3527866	19393
1·5	·0002187	1072	4·0	·3754351	21516
1·6	·0004437	1799	4·1	·3959320	22653
1·7	·0008486	2860	4·2	·4141636	22875
1·8	·0015395 −	4312	4·3	·4301077	22306
1·9	·0026616	6206	4·4	·4438212	21095
2·0	·0044043	8544	4·5	·4554252	19410
2·1	·0070014	11274	4·6	·4650882	17426
2·2	·0107259	14287	4·7	·4730086	15284
2·3	·0158791	17413	4·8	·4794006	13117
2·4	·0227736	20426	4·9	·4844809	11030
			5·0	·4884582	9097
			∞	·5000000	

ON THE MOMENTS OF THE DISTRIBUTION OF SQUARED STANDARD-DEVIATIONS FOR SAMPLES OF N DRAWN FROM AN INDEFINITELY LARGE POPULATION.

By A. E. R. CHURCH, M.A., B.Sc.

(i) Being engaged in an experimental investigation of the moments of the distributions of means and squared standard-deviations (σ^2) of samples of N drawn from a definitely skew population I have had occasion to resort to the formulae for the moments of the distribution of σ^2 for samples of N drawn from an indefinitely large population given by Prof. Al. A. Tchouproff in *Biometrika*, Vol. XII. p. 193 (23) and p. 194 (24) and (25).

Considerable time was spent in an attempt to apply these to the actual distributions used, and finally Prof. Karl Pearson, under whose direction this experimental work is being carried out, came to the definite conclusion that the formula (25) for the 4th moment was in error. I read through the algebraic manipulation which leads up to these numerical results but was unable to follow its final stages where the statement is very exiguous.

Now, previously to Tchouproff, "Student" had given the first and second of these moments in *Biometrika*, Vol. VI. p. 3, although unfortunately, the second moment contained by a printer's failure slight errors corrected, however, in the following line. "Student" also gave the 3rd and 4th moments by the same process in the case of samples from a normal distribution.

In the present paper I have extended "Student's" method to obtain the 3rd and 4th moments for samples from any population so that I might have a revision of Tchouproff's 4th moment by an independent method. I further desired to curtail, if possible, the somewhat extensive use of symbols employed in Prof. Tchouproff's memoir, which discourages the English statistical student.

(ii) Let $\bar{\nu}_t$ be the tth moment of the original indefinitely large population of a variable X, where X is measured from the mean of this population.

Let ${}_2M_t'$ be the tth moment, about $X = 0$, of the distribution of σ^2 of samples of N drawn from this population.

Let ${}_2M_t$ be the tth moment, about its own mean, of this distribution of σ^2.

Then
$$ {}_2M_t' = \Sigma \left[\frac{S(X_1^2)}{N} - \left\{ \frac{S(X_1)}{N} \right\}^2 \right]^t, $$

where S refers to the summation over the N values $X_1, X_2, \ldots X_N$ of the variable X which form one sample, and where, since the original population is indefinitely large, these N values of X may be considered independent or uncorrelated; while by Σ is meant the process of summing for all possible samples of N and dividing by the number of such samples.

To obtain the first four moments ${}_2M_1'$, ${}_2M_2'$, ${}_2M_3'$, ${}_2M_4'$ it is necessary to expand a number of expressions of the form $\{S(X_1)^2\}^p \{S(X_1)\}^{2q}$, where p and q are zero or positive integers less than 8, in series of sums like $S(X_1^{\alpha} X_2^{\beta} \dots X_r^{\lambda})$, where $\alpha, \beta, \dots \lambda$ and r are positive integers or zero, each series of sums being isobaric in the X's of weight m, where $m = 2, 4, 6$, or 8.

Thus I find

$$\{S(X_1^2)\}^2 = S(X_1^4) + 2S(X_1^2X_2^2),$$

$$\{S(X_1^2)\}^3 = S(X_1^6) + 3S(X_1^4X_2^2) + 6S(X_1^2X_2^2X_3^2),$$

$$\{S(X_1^2)\}^4 = S(X_1^8) + 4S(X_1^6X_2^2) + 12S(X_1^4X_2^2X_3^2) + 6S(X_1^4X_2^4) + 24S(X_1^2X_2^2X_3^2X_4^2),$$

$$\{S(X_1)\}^2 = S(X_1^2) + 2S(X_1X_2),$$

$$\{S(X_1)\}^4 = S(X_1^4) + 4S(X_1^3X_2) + 6S(X_1^2X_2^2) + 12S(X_1^2X_2X_3) + 24S(X_1X_2X_3X_4),$$

$$\{S(X_1)\}^6 = S(X_1^6) + 6S(X_1^5X_2) + 15S(X_1^4X_2^2) + 30S(X_1^4X_2X_3) + 20S(X_1^3X_2^3) + 60S(X_1^3X_2^2X_3) + 120S(X_1^3X_2X_3X_4) + 90S(X_1^2X_2^2X_3^2) + 180S(X_1^2X_2^2X_3X_4) + 360S(X_1^2X_2X_3X_4X_5) + 720S(X_1X_2X_3X_4X_5X_6),$$

$$\{S(X_1)\}^8 = S(X_1^8) + 8S(X_1^7X_2) + 28S(X_1^6X_2^2) + 56S(X_1^6X_2X_3) + 56S(X_1^5X_2^3) + 168S(X_1^5X_2^2X_3) + 336S(X_1^5X_2X_3X_4) + 70S(X_1^4X_2^4) + 280S(X_1^4X_2^3X_3) + 420S(X_1^4X_2^2X_3^2) + 840S(X_1^4X_2^2X_3X_4) + 1680S(X_1^4X_2X_3X_4X_5) + 560S(X_1^3X_2^3X_3^2) + 1120S(X_1^3X_2^3X_3X_4) + 1680S(X_1^3X_2^2X_3^2X_4) + 3360S(X_1^3X_2^2X_3X_4X_5) + 6720S(X_1^3X_2X_3X_4X_5X_6) + 2520S(X_1^2X_2^2X_3^2X_4^2) + 5040S(X_1^2X_2^2X_3^2X_4X_5) + 10080S(X_1^2X_2^2X_3X_4X_5X_6) + 20160S(X_1^2X_2X_3X_4X_5X_6X_7) + 40320S(X_1X_2X_3X_4X_5X_6X_7X_8),$$

$$\{S(X_1^2)\}^2\{S(X_1)\}^2 = S(X_1^6) + 2S(X_1^5X_2) + 3S(X_1^4X_2^2) + 2S(X_1^4X_2X_3) + 4S(X_1^3X_2^3) + 4S(X_1^3X_2^2X_3) + 6S(X_1^2X_2^2X_3^2) + 4S(X_1^2X_2^2X_3X_4),$$

$$\{S(X_1^2)\}^2\{S(X_1)\}^4 = S(X_1^8) + 4S(X_1^7X_2) + 8S(X_1^6X_2^2) + 12S(X_1^6X_2X_3) + 12S(X_1^5X_2^3) + 20S(X_1^5X_2^2X_3) + 24S(X_1^5X_2X_3X_4) + 14S(X_1^4X_2^4) + 32S(X_1^4X_2^3X_3) + 28S(X_1^4X_2^2X_3^2) + 36S(X_1^4X_2^2X_3X_4) + 24S(X_1^4X_2X_3X_4X_5) + 40S(X_1^3X_2^3X_3^2) + 56S(X_1^3X_2^3X_3^2X_4) + 48S(X_1^3X_2^2X_3^2X_4) + 48S(X_1^3X_2^2X_3X_4X_5) + 72S(X_1^3X_2^2X_3^2X_4^2) + 72S(X_1^3X_2^2X_3^2X_4X_5) + 48S(X_1^3X_2^2X_3X_4X_5X_6),$$

$$\{S(X_1^2)\}^3\{S(X_1)\}^2 = S(X_1^8) + 2S(X_1^7X_2) + 4S(X_1^6X_2^2) + 2S(X_1^6X_2X_3) + 6S(X_1^6X_3^2) + 6S(X_1^5X_2^2X_3) + 6S(X_1^4X_2^4) + 6S(X_1^4X_2^2X_3) + 12S(X_1^4X_2^2X_3^2) + 6S(X_1^4X_2^2X_3X_4) + 12S(X_1^3X_2^3X_3^2) + 12S(X_1^3X_2^2X_3^2X_4) + 24S(X_1^2X_2^2X_3^2X_4^2) + 12S(X_1^2X_2^2X_3^2X_4X_5),$$

$$\{S(X_1^2)\}\{S(X_1)\}^2 = S(X_1^4) + 2S(X_1^3X_2) + 2S(X_1^2X_2^2) + 2S(X_1^2X_2X_3),$$

$$\begin{aligned}\{S(X_1^2)\}\{S(X_1)\}^4 = {} & S(X_1^6) + 4S(X_1^5X_2) + 7S(X_1^4X_2^2) + 12S(X_1^4X_2X_3) \\ & + 8S(X_1^3X_2^3) + 16S(X_1^3X_2^2X_3) + 24S(X_1^3X_2X_3X_4) \\ & + 18S(X_1^2X_2^2X_3^2) + 24S(X_1^2X_2^2X_3X_4) + 24S(X_1^2X_2X_3X_4X_5),\end{aligned}$$

$$\begin{aligned}\{S(X_1^2)\}\{S(X_1)\}^6 = {} & S(X_1^8) + 6S(X_1^7X_2) + 16S(X_1^6X_2^2) + 30S(X_1^6X_2X_3) \\ & + 26S(X_1^5X_2^3) + 66S(X_1^5X_2^2X_3) + 120S(X_1^5X_2X_3X_4) \\ & + 30S(X_1^4X_2^4) + 90S(X_1^4X_2^3X_3) + 120S(X_1^4X_2^2X_3^2) \\ & + 210S(X_1^4X_2^2X_3X_4) + 360S(X_1^4X_2X_3X_4X_5) + 140S(X_1^3X_2^3X_3^2) \\ & + 300S(X_1^3X_2^2X_3^2X_4) + 240S(X_1^3X_2^3X_3X_4) \\ & + 480S(X_1^3X_2^2X_3X_4X_5) + 720S(X_1^3X_2X_3X_4X_5X_6) \\ & + 360S(X_1^2X_2^2X_3^2X_4^2) + 540S(X_1^2X_2^2X_3^2X_4X_5) \\ & + 720S(X_1^2X_2^2X_3X_4X_5X_6) + 720S(X_1^2X_2X_3X_4X_5X_6X_7).\end{aligned}$$

Now if $\alpha, \beta, \gamma, \ldots$ be all different, $S(X_1^\alpha X_2^\beta \ldots X_r^\lambda)$ has $N(N-1)\ldots(N-r+1)$ terms in its expansion; but if p indices be alike of one kind α, and q be alike of another kind β, etc., it has $\dfrac{N(N-1)\ldots(N-r+1)}{p!\,q!\ldots}$ terms in its expansion:

$$\therefore\ \Sigma S(X_1^\alpha X_2^\beta \ldots X_r^\lambda) = N(N-1)\ldots(N-r+1)\,\Sigma'(X_1^\alpha X_2^\beta \ldots X_r^\lambda),$$

or
$$\frac{N(N-1)\ldots(N-r+1)}{p!\,q!\ldots}\,\Sigma'\{X_1^\alpha X_2^\alpha \ldots X_p^\alpha X_{p+1}^\beta \ldots X_{p+q}^\beta \ldots\},$$

as the case may be, where Σ' refers to the summation over all values of the variate X in the original population, divided by the number of such values of course.

Now since $X_1, X_2, \ldots X_N$ are independent or uncorrelated we have therefore

$$\Sigma S(X_1^\alpha X_2^\beta \ldots X_r^\lambda) = N(N-1)\ldots(N-r+1)\,\Sigma' X_1^\alpha \Sigma' X_2^\beta \ldots \Sigma' X_r^\lambda$$

or
$$\frac{N(N-1)\ldots(N-r+1)}{p!\,q!\ldots}(\Sigma' X_1^\alpha)^p(\Sigma' X_2^\beta)^q \ldots,$$

that is
$$\Sigma S(X_1^\alpha X_2^\beta \ldots X_r^\lambda) = N(N-1)\ldots(N-r+1)\,\bar\nu_\alpha \bar\nu_\beta \ldots \bar\nu_\lambda,$$

or
$$\frac{N(N-1)\ldots(N-r+1)}{p!\,q!\ldots}\bar\nu_\alpha{}^p \bar\nu_\beta{}^q \ldots;$$

but since these moments of the original distribution are about its mean $X = 0$, we have $\bar\nu_1 = 0$, therefore if any of $\alpha, \beta, \ldots \lambda$ are unity, $\Sigma S(X_1^\alpha X_2^\beta \ldots X_r^\lambda) = 0$.

Therefore, using the above, we have

$${}_2M_1' = \Sigma\left[\frac{S(X_1^2)}{N} - \frac{\{S(X_1)\}^2}{N^2}\right] = \bar\nu_2\left(1 - \frac{1}{N}\right),$$

also
$${}_2M_2' = \Sigma\left[\frac{\{S(X_1^2)\}^2}{N^2} - \frac{2\{S(X_1^2)\}\{S(X_1)\}^2}{N^3} + \frac{\{S(X_1)\}^4}{N^4}\right],$$

i.e.
$${}_2M_2' = \left(1 - \frac{1}{N}\right)\left[\frac{1}{N}\left(1 - \frac{1}{N}\right)\bar\nu_4 + \left(1 - \frac{2}{N} + \frac{3}{N^2}\right)\bar\nu_2^2\right].$$

Further

$$ {}_2M_3' = \Sigma\left[\frac{\{S(X_1^2)\}^3}{N^3} - \frac{3\{S(X_1^2)\}^2\{S(X_1)\}^2}{N^4} + \frac{3\{S(X_1^2)\}\{S(X_1)\}^4}{N^5} - \frac{\{S(X_1)\}^6}{N^6}\right], $$

i.e. $$ {}_2M_3' = \bar{\nu}_2^3 + \frac{1}{N}[3\bar{\nu}_4\bar{\nu}_2 - 6\bar{\nu}_2^3] + \frac{1}{N^2}[\bar{\nu}_6 - 12\bar{\nu}_4\bar{\nu}_2 - 6\bar{\nu}_3^2 + 20\bar{\nu}_2^3] $$

$$ - \frac{1}{N^3}[3\bar{\nu}_6 - 30\bar{\nu}_4\bar{\nu}_2 - 18\bar{\nu}_3^2 + 48\bar{\nu}_2^3] + \frac{1}{N^4}[3\bar{\nu}_6 - 36\bar{\nu}_4\bar{\nu}_2 - 22\bar{\nu}_3^2 + 63\bar{\nu}_2^3] $$

$$ - \frac{1}{N^5}[\bar{\nu}_6 - 15\bar{\nu}_4\bar{\nu}_2 - 10\bar{\nu}_3^2 + 30\bar{\nu}_2^3]. $$

Again

$$ {}_2M_4' = \Sigma\left[\frac{\{S(X_1^2)\}^4}{N^4} - \frac{4\{S(X_1^2)\}^3\{S(X_1)\}^2}{N^5} + \frac{6\{S(X_1^2)\}^2\{S(X_1)\}^4}{N^6} - \frac{4\{S(X_1^2)\}\{S(X_1)\}^6}{N^7} + \frac{\{S(X_1)\}^8}{N^8}\right], $$

i.e. $$ {}_2M_4' = \bar{\nu}_2^4 + \frac{1}{N}[6\bar{\nu}_4\bar{\nu}_2^2 - 10\bar{\nu}_2^4] + \frac{1}{N^2}[4\bar{\nu}_6\bar{\nu}_2 + 3\bar{\nu}_4^2 - 42\bar{\nu}_4\bar{\nu}_2^2 - 24\bar{\nu}_3^2\bar{\nu}_2 + 53\bar{\nu}_2^4] $$

$$ + \frac{1}{N^3}[\bar{\nu}_8 - 20\bar{\nu}_6\bar{\nu}_2 - 24\bar{\nu}_5\bar{\nu}_3 - 15\bar{\nu}_4^2 + 180\bar{\nu}_4\bar{\nu}_2^2 + 192\bar{\nu}_3^2\bar{\nu}_2 - 218\bar{\nu}_2^4] $$

$$ - \frac{1}{N^4}[4\bar{\nu}_8 - 64\bar{\nu}_6\bar{\nu}_2 - 96\bar{\nu}_5\bar{\nu}_3 - 54\bar{\nu}_4^2 + 576\bar{\nu}_4\bar{\nu}_2^2 + 688\bar{\nu}_3^2\bar{\nu}_2 - 687\bar{\nu}_2^4] $$

$$ + \frac{1}{N^5}[6\bar{\nu}_8 - 112\bar{\nu}_6\bar{\nu}_2 - 176\bar{\nu}_5\bar{\nu}_3 - 102\bar{\nu}_4^2 + 1122\bar{\nu}_4\bar{\nu}_2^2 + 1360\bar{\nu}_3^2\bar{\nu}_2 - 1398\bar{\nu}_2^4] $$

$$ - \frac{1}{N^6}[4\bar{\nu}_8 - 92\bar{\nu}_6\bar{\nu}_2 - 160\bar{\nu}_5\bar{\nu}_3 - 95\bar{\nu}_4^2 + 1110\bar{\nu}_4\bar{\nu}_2^2 + 1400\bar{\nu}_3^2\bar{\nu}_2 - 1515\bar{\nu}_2^4] $$

$$ + \frac{1}{N^7}[\bar{\nu}_8 - 28\bar{\nu}_6\bar{\nu}_2 - 56\bar{\nu}_5\bar{\nu}_3 - 35\bar{\nu}_4^2 + 420\bar{\nu}_4\bar{\nu}_2^2 + 560\bar{\nu}_3^2\bar{\nu}_2 - 630\bar{\nu}_2^4]. $$

Now changing the origin from the mean of the original population to the mean of this distribution of σ^2 of samples of N we have for the equations of transformation the usual formulae:

$$ {}_2M_2 = {}_2M_2' - ({}_2M_1')^2, $$

$$ {}_2M_3 = {}_2M_3' - 3\,.\,{}_2M_2\,.\,{}_2M_1' - ({}_2M_1')^3, $$

$$ {}_2M_4 = {}_2M_4' - 4\,.\,{}_2M_3\,.\,{}_2M_1' - 6\,.\,{}_2M_2\,.\,({}_2M_1')^2 - ({}_2M_1')^4, $$

which give

$$ {}_2M_2 = \frac{1}{N}\left(1 - \frac{1}{N}\right)\left[\left(1 - \frac{1}{N}\right)\bar{\nu}_4 - \left(1 - \frac{3}{N}\right)\bar{\nu}_2^2\right] $$

$$ = \frac{1}{N}[\bar{\nu}_4 - \bar{\nu}_2^2] - \frac{2}{N^2}[\bar{\nu}_4 - 2\bar{\nu}_2^2] + \frac{1}{N^3}[\bar{\nu}_4 - 3\bar{\nu}_2^2] \quad \text{.....(i)}, $$

$$ {}_2M_3 = \frac{1}{N^2}[\bar{\nu}_6 - 3\bar{\nu}_4\bar{\nu}_2 - 6\bar{\nu}_3^2 + 2\bar{\nu}_2^3] - \frac{1}{N^3}[3\bar{\nu}_6 - 21\bar{\nu}_4\bar{\nu}_2 - 18\bar{\nu}_3^2 + 26\bar{\nu}_2^3] $$

$$ + \frac{1}{N^4}[3\bar{\nu}_6 - 33\bar{\nu}_4\bar{\nu}_2 - 22\bar{\nu}_3^2 + 54\bar{\nu}_2^3] - \frac{1}{N^5}[\bar{\nu}_6 - 15\bar{\nu}_4\bar{\nu}_2 - 10\bar{\nu}_3^2 + 30\bar{\nu}_2^3] \quad \text{(ii)}, $$

and

$$
\begin{aligned}
{}_2M_4 = {} & \frac{3}{N^2}[\bar{\nu}_4 - \bar{\nu}_2^2]^2 + \frac{1}{N^3}[\bar{\nu}_8 - 4\bar{\nu}_6\bar{\nu}_2 - 24\bar{\nu}_5\bar{\nu}_3 - 15\bar{\nu}_4^2 + 48\bar{\nu}_4\bar{\nu}_2^2 + 96\bar{\nu}_3^2\bar{\nu}_2 - 30\bar{\nu}_2^4] \\
& - \frac{1}{N^4}[4\bar{\nu}_8 - 40\bar{\nu}_6\bar{\nu}_2 - 96\bar{\nu}_5\bar{\nu}_3 - 54\bar{\nu}_4^2 + 336\bar{\nu}_4\bar{\nu}_2^2 + 528\bar{\nu}_3^2\bar{\nu}_2 - 306\bar{\nu}_2^4] \\
& + \frac{1}{N^5}[6\bar{\nu}_8 - 96\bar{\nu}_6\bar{\nu}_2 - 176\bar{\nu}_5\bar{\nu}_3 - 102\bar{\nu}_4^2 + 924\bar{\nu}_4\bar{\nu}_2^2 + 1232\bar{\nu}_3^2\bar{\nu}_2 - 1044\bar{\nu}_2^4] \\
& - \frac{1}{N^6}[4\bar{\nu}_8 - 88\bar{\nu}_6\bar{\nu}_2 - 160\bar{\nu}_5\bar{\nu}_3 - 95\bar{\nu}_4^2 + 1050\bar{\nu}_4\bar{\nu}_2^2 + 1360\bar{\nu}_3^2\bar{\nu}_2 - 1395\bar{\nu}_2^4] \\
& + \frac{1}{N^7}[\bar{\nu}_8 - 28\bar{\nu}_6\bar{\nu}_2 - 56\bar{\nu}_5\bar{\nu}_3 - 35\bar{\nu}_4^2 + 420\bar{\nu}_4\bar{\nu}_2^2 + 560\bar{\nu}_3^2\bar{\nu}_2 - 630\bar{\nu}_2^4] \quad \ldots(\text{iii}).
\end{aligned}
$$

A comparison between the formulae given above for ${}_2M_4'$ and ${}_2M_4$ with Tchouproff's formula (25) appears to indicate some algebraical slip in his substitution of the value of his $W^{(4)}_{2(N)}$ of equation (21) of p. 193 in the third equation of his (10) p. 187 to get his (25) of p. 194. For his coefficients in the terms of (25) involving $1/N^6$ are exactly those of $1/N^2$ in his $W^{(4)}_{2(N)}$, or of the term in $1/N^6$ in my ${}_2M_4'$, and not those of this fourth moment, our ${}_2M_4$, when transferred to the mean σ^2 of samples.

The method, due to "Student," employed here is undoubtedly more direct and concise in statement than that employed by Prof. Tchouproff; further it presents fewer algebraical difficulties and the algebra has the advantage of being easily followed and checked at each stage.

It is difficult to estimate at present the real value of these high moment formulae in actual practice, but it is fairly clear that, if they be used to predict the results of any actual sampling, that sampling must be strictly in accordance with the assumptions involved and must be very carefully carried out in experimental work before any reasonable agreement can be obtained.

In conclusion I must express my thanks to Prof. K. Pearson for his continual kindly advice and criticism.

THE CORPUSCLE PROBLEM.

A MATHEMATICAL STUDY OF A BIOMETRIC PROBLEM.

By S. D. WICKSELL, *Lund.*

1. In several organs of the human body, as well as that of many animals, a microscopic and sometimes also a macroscopic inspection reveals a number of small, more or less regularly formed corpuscles embedded in the tissue of the organ. Such corpuscles are for instance the Germ Centres or secondary follicles (Hellman) found in lymphoidal organs—as the lymphatic glands, the tonsils or the spleen,—the so-called *Hassal corpuscles* in the thymus, the *Langerhans insulae* in the pancreas, etc. Generally these bodies vary considerably in size within the same organ, and in different organs or different individuals the number, as well as the size distribution of the corpuscles will be found to be different.

In anatomy as in other biological disciplines the last few decades have seen an enormous increase in the demand for exact and detailed information. The modern anatomist often devotes the labour and toil of many years to the study of one special organ, patiently collecting and measuring extensive material from animals or *post mortem* subjects. The more extensive the material the more difficult it will be, however, to perform for each individual all the detailed observations required; the investigator will have to look for simplifications and short-cut methods of observation, as far as this is possible without making the resulting information insufficient. In studying the number and size distributions of the above-mentioned corpuscles, the anatomists have, then, run up against a mathematical problem, which is, even in the simpler cases, beyond the powers of persons unacquainted with higher mathematics. This problem, which I call the "corpuscle problem," is, as far as I know, as yet unsolved, although several anatomists have tried to reach a solution along more or less purely empirical lines*. At the request of Professor T. Hellman, of Lund, I have therefore taken the question up from the mathematico-statistical point of view. My interest in the matter has subsequently been greatly increased by the rather curious fact, that the problem is in essential features identical with an important problem in astronomical statistics.

2. In order to make the matter more concrete I shall begin with a short description of the observations performed by Hellman regarding the so-called secondary follicles (Germ Centres) in the lymphatic tissue of the spleen.

* See for instance J. A. Hammar, "Methode, die Menge der Rinde und des Marks des Thymus, sowie die Anzahl und die Grösse der Hassal'schen Körper zahlenmässig festzustellen" (*Zeitschrift f. angewandte Anatomie und Konstitutionslehre*, Bd. I, Heft 4/5, 1914); J. A. Hammar and T. Hellman, "Ein Fall von Tyreoaplasie unter Berücksichtigung der innersekretorischen und lymphoiden Organe: Abteilung II: Die lymphoiden Organe," von T. Hellman (*Zeitschrift f. die gesammte Anatomie*, Zweite Abteilung, Bd. VIII, Heft 4, 1922).

The Germ Centre, or rather the secondary follicle in the Malpighian corpuscle, is a small, approximately spherical, body embedded in the tissue of the spleen. The number of such corpuscles in different spleens is very variable, ranging from a very few or none at all up to hundreds of thousands per organ. The sizes are also very variable, the same organ often containing all sizes from a small fraction of a millimetre up to one millimetre in diameter.

The observations are arranged in the following way: The weight of each spleen being ascertained, and fixation and preparation being properly attended to, microtome sections of 0·018 mm. in thickness are taken from different parts of the organ. By the aid of a projection apparatus a picture of each section is obtained in 18-fold magnification on a paper, and the contours (generally circular, sometimes oval) of the follicles are traced on the paper in lead pencil. On the drawings thus obtained the diameter of every corpuscle contour is measured to the nearest millimetre; when the image is not circular the largest diameter and also the diameter at right angles to the largest are measured, the arithmetic mean of the two measures being taken as value of the diameter. When the diameter is below 1·5 mm. in this magnification the measures are rejected on account of difficulties of identification. From the protocols kept of these measures the distribution of the diameters of the images in the section is obtained for classes of one mm. breadth (1·5—2·5, 2·5—3·5, etc.).

The outline of the preparation also being marked out in the drawing the parts of the paper representing the spleen parenchyma may be cut out as well as the images of the follicles. As a carefully selected homogeneous paper is always used the areas of the sections may be determined by weighing the drawings. The weights of the drawings before and after the images of the follicles have been cut away will also give sample values of the ratio of the total volume of the follicles to the total volume of the spleen. As will be seen later on it is necessary to know either the area of the section or this ratio in order to determine the total number of follicles in the organ. Of course, when several sections have been taken and measured from different parts of the same organ, they are put together and regarded as one single section of correspondingly greater area. In the following we will regard this area as great enough for the purposes of generalization. Thus the question of the errors of sampling will not be taken up in this connection.

It is furthermore clear that the diameters of the images in the drawings are generally not the true diameters of the follicles. Only when the section chances to pass through the centre of a follicle will this be the case. As a rule the centres of the follicles will lie more or less distant from the plane of the section. The more distant the centre the smaller will the image diameter be, in comparison to the actual diameter. As the distance of the centre is generally unknown we see that the diameters measured in the drawings are "apparent diameters" in the same sense as the projection on the heavens of the motion of a star or the distance of the components of a double star is called apparent, the inclination to the plane of projection of the true motion or line joining the components being unknown. Indeed, the difficulty of the anatomist at his microscope is essentially the same as

the difficulty of the astronomer at his telescope; he only sees the projection of things. The statistical problem met with is also in both cases similar. The "true" character being a function of the apparent character and an unknown variable, and making a plausible assumption as to the distribution of this variable, the problem is to express the distribution of the true character in terms of the observed distribution of the apparent character. It will be seen later on that there is also one problem in stellar statistics, which is in essential details identical to the anatomic corpuscle problem, namely the problem of determining the distribution in space of the stars in a stellar cluster, from the projected distribution on the heavens. Hence some of the methods here worked out for anatomic applications may be of value also for the astronomers.

Evidently the problem arising in connection with the secondary follicles of the spleen is the following: To find the distribution of the actual ("true") diameters of the follicles from the distribution of the diameters of the images (the "apparent" diameters) in the microtome sections. That these two distributions are generally different is clear from the following two reasons: (1) The apparent diameters are as a rule smaller than the true diameters, and (2) a random section will contain a relatively greater number of actually large follicles, than of small ones, because the former more frequently will come within reach of the section plane. The two causes will, however, work in opposite directions, thus having a chance to balance each other. It will be seen in the following that under certain conditions an exact balance takes place, making the relative distributions of the true and the apparent diameters equal.

We must mention finally that in taking a series of successive sections instead of only one or a few sample sections, the true diameters may be directly observed. In this respect the position of the microscopist is more favourable than that of his colleague at the telescope. But as the measuring of one single section is generally very laborious, and as the material studied embraces several hundred spleens, it has been found impossible to use this method, except in special cases to obtain a control and empirical test of the theoretical reductions. These reductions naturally involve the assumption that the corpuscles are approximately spherical in form, and that they are fairly uniformly distributed throughout the whole organ. Series of successive sections actually measured show these conditions to be sufficiently fulfilled in the case of the secondary follicles of the spleen. Attempts have been made to use the serial sections for obtaining a set of purely empirical equations of reduction. These trials have, however, shown that the theoretical equations are more to be relied upon.

I now pass on to a more general formulation and solution of the "corpuscle problem."

3. *The Corpuscle Problem.* In an opaque body there are suspended a large number of spherical corpuscles of different sizes, the density of the centres and the distribution of the sizes being the same in all parts of the body. This body is split in two by a plane section. To express the distribution of the diameters of the

corpuscles in terms of the distribution of the diameters of the circular contours found in the plane section.

Solution. Let the number of centres per unit volume be N and let $F(r)\,dr$ be the relative frequency of corpuscles with the diameter in the interval r to $r+dr$. We now have

$$\int_0^R F(r)\,dr = 1 \quad \text{.....(1).}$$

R being the upper limit to the diameters. Furthermore the number of corpuscles of diameter r to $r+dr$ coming within reach of an arbitrary plane is

$$N r\,F(r)\,dr$$

per unit area in the plane. Hence, r_0 being the mean diameter of the corpuscles· Nr_0 will be the total number of corpuscle images to the unit area in the section plane. Putting

$$Nr_0 f(r) = N r\,F(r) \quad \text{.....(2).}$$

it is seen that $f(r)\,dr$ is the relative frequency of corpuscles cut by the plane. having the actual diameter between r and $r+dr$. We thus have

$$f(r) = \frac{rF(r)}{r_0} \quad \text{.....(3).}$$

The probability that a corpuscle with the diameter r, falling within the reach of the plane section, has its centre at the distance y to $y+dy$ from this plane, is evidently equal to

$$\frac{2dy}{r}.$$

Hence the relative number of corpuscles of diameter r to $r+dr$ and with their centres at the distance y to $y+dy\,(y<r)$ from the plane, is equal to

$$\frac{2}{r} f(r)\,dy\,dr.$$

But the "apparent" diameter of the circular image cut out of a corpuscle by the plane being denoted by x, we have

$$x^2 = r^2 - 4y^2,$$

and it is seen that the relative number of corpuscles having the actual diameter r to $r+dr$ and the apparent diameter x to $x+dx$ in the section plane is

$$\frac{2}{r} f(r)\,dr\,\frac{dy}{dx}\,dx = \frac{1}{r} f(r)\,dr\,\frac{x\,dx}{\sqrt{r^2-x^2}} = \frac{1}{r_0} F(r)\,dr\,\frac{x\,dx}{\sqrt{r^2-x^2}} \quad \text{.....(4).}$$

Hence, $\phi(x)$ being the relative frequency function of the apparent diameters in the plane section, we have

$$\phi(x) = \frac{x}{r_0}\int_x^R F(r)\,\frac{dr}{\sqrt{r^2-x^2}} \quad \text{.....(5).}$$

As $\phi(x)$ is here the function supposed to be known, this is an integral equation to determine the sought function $F(r)$.

This integral equation may easily be transformed to the same form as the well-known integral equation of Abel*, and thus the solution is readily obtained. It will be still more convenient, however, to apply the solution given by von Zeipel of the following equation, occurring in the reduction of the observations of globular clusters†,

$$\psi(x) = 2\int_x^R \chi(r)\frac{r\,dr}{\sqrt{r^2-x^2}} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(6).$$

For this equation von Zeipel has obtained the solution

$$\chi(r) = -\frac{1}{\pi}\int_r^R \frac{\psi'(x)\,dx}{\sqrt{x^2-r^2}} = \frac{1}{\pi}\int_0^{\sqrt{R^2-r^2}} P(\sqrt{l^2+r^2})\,dl \quad \ldots\ldots\ldots\ldots(7),$$

where
$$P(x) = -\frac{1}{x}\frac{d\psi(x)}{dx}.$$

The solution of (5) consequently assumes the attractive form

$$F(r) = -\frac{2rr_0}{\pi}\int_r^R \frac{d\left(\frac{\phi(x)}{x}\right)}{dx}\frac{dx}{\sqrt{x^2-r^2}} = \frac{2rr_0}{\pi}\int_0^{\sqrt{R^2-r^2}} P(\sqrt{l^2+r^2})\,dl \quad \ldots(8),$$

where
$$P(x) = -\frac{1}{x}\frac{d\left(\frac{\phi(x)}{x}\right)}{dx} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(9).$$

4. This form of solution is also specially adapted to the case, in which we do not know an analytical expression for $\phi(x)$, but only have a series of numerical values of this function, which is what takes place both in the astronomical problem as well as in the anatomical problem here in question.

The following method will certainly be sufficiently accurate for the anatomical applications: The observations in the sections having been arranged in classes, the relative frequency-histogram of the apparent diameters may be constructed. Smoothing this by a continuous curve (a free-hand curve will suffice) we may read off the ordinates $\phi(i)$; $(i = 0, 1, 2, \ldots R)$. Thence we obtain the series $\phi(1) : 1$; $\phi(2) : 2$; $\phi(3) : 3$, etc., and finally the differential quotients of $\phi(x) : x$ may be obtained by the aid of the well-known central difference formula

$$f'(x) = \Delta_0^1 f(x) - \tfrac{1}{6}\Delta_0^3 f(x) + \tfrac{1}{30}\Delta_0^5 f(x) + \ldots.$$

From these we easily obtain the values of $P(i)$, and the curve $y = P(x)$ may be plotted, so that we may read off the values of

$$P(\sqrt{i^2+r^2}); \quad (i = 0, 1, 2 \ldots R).$$

* See for instance Bôcher, *An Introduction to the Study of Integral Equations*, pp. 6—11 (Cambridge Tracts in Mathematics and Mathematical Physics), 1909.

† H. von Zeipel, "Catalogue de l'amas globulaire Messier 3" (*Annales de l'Observatoire de Paris*, Tome XXV, 1906, p. 30); "Recherches sur la constitution des amas globulaires" (*Kungl. Svenska Vetenskapsakademiens Handlingar*, Bd. LI, No. 5, 1913, p. 9). H. von Zeipel und J. Lindgren, "Photometrische Untersuchungen der Sterngruppe Messier 37" (*Kungl. Svenska Vetenskapsakademiens Handlingar*, Bd. LXI, No. 15, 1921, pp. 109—110).

Using the approximate quadrature formula *

$$\frac{2r}{\pi}\int_0^{\sqrt{R^2-r^2}} P(\sqrt{l^2+r^2})\,dl = \frac{2r}{\pi}\left[\sum_{i=0}^{\sqrt{R^2-r^2}} P(\sqrt{i^2+r^2}) - \tfrac{1}{2}P(r)\right] = A(r) \quad \ldots(10),$$

we have

$$F(r) = r_0 A(r).$$

Since

$$\int_0^R F(r)\,dr = 1,$$

we evidently must have approximately $\frac{1}{r_0} = \sum_{r=0}^{R} A(r)$. The mean diameter r_0 being found in this manner, a graph of the sought function $F(r)$ is easily constructed.

To determine the total number N of follicles per unit volume we may use the results obtained by weighing the section-drawings. These weighings give, as mentioned in § 2, the volume of the follicles in percentages of the spleen tissue. Denoting this by $100\,V$, and the mean volume of the follicles by v_0, we evidently have

$$N = V : v_0.$$

The mean volume v_0 is obtained from the equation

$$v_0 = \frac{\pi}{6}\sum_{r=0}^{R} r^3 F(r).$$

The weighings of the drawings may, however, also be used to determine the area of the observed sections. This area being denoted by Y, and the number of follicles represented in the observed sections being denoted by n, we evidently have

$$n = N r_0 Y,$$

from which equation N is readily found. It may be presumed that this method is subject to a smaller mean error than the preceding method.

5. Another solution of (5), which is perhaps more to the taste of the trained statistician, is obtained by the aid of the method of moments.

The quantities v_n' and N_n', defined by the equations

$$v_n' = \int_0^R x^n \phi(x)\,dx; \quad N_n' = \int_0^R r^n F(r)\,dr,$$

are known as the moments about the origin of the frequency functions $\phi(x)$ and $F(r)$. As is well known from the papers of Professor Karl Pearson, it is generally possible to construct the frequency function, when the numerical values of a sufficient number of the moments have been obtained. As a rule the first four moments, in exceptional cases the first six moments, are sufficient.

Now, the moments v_n' of $\phi(x)$ may be numerically determined from the observed distribution of the apparent diameters. The mechanical quadrature formulae for this determination need not be entered into here, as they are familiar to any statistician. Hence the solution of (5) is reduced to the following problem: To express the moments N_n' of the sought function $F(r)$ in terms of the moments v_n'.

* von Zeipel has used this formula in his cluster reductions, where the accuracy required is much greater than in the case of the anatomical applications.

Indeed, one such relation may be written out at once. As the plane section contains Nr_0 corpuscle images per unit area (see § 3), it is clear that $N.N_1' \frac{\pi}{4} v_2'$ (the mean diameter r_0 being here denoted by N_1') is the area of the images per unit area in the sections. Furthermore $N \frac{\pi}{6} N_3'$ is the volume of the corpuscles per unit volume of the tissue. As these two quantities are evidently to be put equal to each other, we have

$$N_3' = \tfrac{3}{2} N_1' v_2'.$$

A general solution is obtained as follows: We have, putting $r_0 = N_1'$ in (5),

$$\phi(x) = \frac{x}{N_1'} \int_x^R F(r) \frac{dr}{\sqrt{r^2 - x^2}}.$$

Hence

$$v_n' = \frac{1}{N_1'} \int_0^R x^{n+1}\, dx \int_x^R F(r) \frac{dr}{\sqrt{r^2 - x^2}}.$$

Using Dirichlet's formula* this reduces to

$$v_n' = \frac{1}{N_1'} \int_0^R F(r)\, dr \int_0^r \frac{x^{n+1}}{\sqrt{r^2 - x^2}} dx.$$

Putting

$$y_n = \int_0^r \frac{x^{n+1}}{\sqrt{r^2 - x^2}} dx,$$

it follows from well-known theorems in the theory of integration, that we have

$$y_0 = r;\ y_1 = \frac{\pi}{4} r^2;\ y_2 = \tfrac{2}{3} r^3;\ y_3 = \frac{3\pi}{16} r^4;\ y_4 = \tfrac{8}{15} r^5;\ y_5 = \frac{5\pi}{32} r^6;\ y_6 = \tfrac{16}{35} r^7.$$

Hence we finally get

$$\left.\begin{aligned} &N_1' = N_1' v_0' \text{ (i.e., as should be, } v_0' = 1), \\ &N_2' = \frac{4}{\pi} N_1' v_1';\ N_3' = \tfrac{3}{2} N_1' v_2';\ N_4' = \frac{16}{3\pi} N_1' v_3'; \\ &N_5' = \tfrac{15}{8} N_1' v_4';\ N_6' = \frac{32}{5\pi} N_1' v_5';\ N_7' = \tfrac{35}{16} N_1' v_6'. \end{aligned}\right\} \quad \ldots\ldots\ldots\ldots (11)$$

It thus only remains to determine N_1', and the problem to express the "true moments" in terms of the "apparent moments" is solved.

In order to determine N_1' we may for instance make use of the proposition: *The arithmetic mean of the true diameters is equal to the harmonic mean of the apparent diameters multiplied by* $\pi/2$. The proof of this proposition is evidently

$$N_1 \int_0^R \phi(x) \frac{dx}{x} = \int_0^R dx \int_x^R F(r) \frac{dr}{\sqrt{r^2 - x^2}} = \int_0^R F(r)\, dr \int_0^r \frac{dx}{\sqrt{r^2 - x^2}} = \frac{\pi}{2} \int_0^R F(r)\, dr = \frac{\pi}{2}.$$

The moments of the sought function $F(r)$ thus being obtained, the construction of the function only depends on a suitable choice of an analytical expression. I

* As to the applicability of Dirichlet's formula when the integrand becomes infinite at the limits of integration see Bôcher, *loc. cit.* p. 4.

believe that Pearson's theory of frequency curves and the auxiliary tables given by him for the construction work, give ample means to meet any requirements in this respect.

It may be added, that the general relations between the moments N' and ν' may be expressed as follows:

$$N_1'\nu_n' = \begin{cases} N'_{n+1}\dfrac{1\,.\,3\,.\,5\,\ldots\,n}{2\,.\,4\,.\,6\,\ldots\,(n+1)}\cdot\dfrac{\pi}{2}, & \text{when } n \text{ is odd,} \\ N'_{n+1}\dfrac{2\,.\,4\,.\,6\,\ldots\,n}{3\,.\,5\,.\,7\,\ldots\,(n+1)}, & \text{when } n \text{ is even.} \end{cases}$$

6. In § 2 it was pointed out that under certain conditions the true distribution $F(r)$ and the apparent distribution $\phi(x)$ become identical. In order to show this we put in (5) $\phi(x) \equiv F(x)$, thus getting the equation of condition

$$F(x) = \frac{x}{r_0}\int_x^R F(r)\,\frac{dr}{\sqrt{r^2 - x^2}}.$$

Putting $F(x) : x = F_1(x)$, we get instead

$$F_1(x) = \frac{1}{r_0}\int_x^R F_1(r)\,\frac{r\,dr}{\sqrt{r^2 - x^2}} = \frac{1}{r_0}\int_0^{\sqrt{R^2 - x^2}} F_1\left(\sqrt{x^2 + y^2}\right) dy.$$

It is seen at once that this equation is satisfied by the function

$$F_1(x) = ke^{-\frac{x^2}{2\sigma^2}},$$

if, as in this case must be done, we put $R = \infty$. Hence as we must also have

$$\int_0^\infty F(r)\,dr = 1,$$

we get

$$F(r) = \frac{1}{\sigma^2} re^{-\frac{r^2}{2\sigma^2}},$$

and

$$r_0 = \sigma\sqrt{\frac{\pi}{2}}.$$

Our conclusion is that if the true distribution has the form

$$F(r) = \frac{r}{\sigma^2} e^{-\frac{r^2}{2\sigma^2}},$$

then the apparent distribution has the same form, or

$$\phi(x) = \frac{x}{\sigma^2} e^{-\frac{x^2}{2\sigma^2}}.$$

That this result is true also in the inverse sense may be shown by the aid of the solution (8) of (5).

It might at first seem as if the same property is at hand when we have the more general form

$$\phi(x) = \sum_1^n k_n \frac{x}{\sigma_n^2} e^{-\frac{x^2}{2\sigma_n^2}}.$$

It is easily found, however, that in this case we get

$$F(r) = \sum_1^n k_n \frac{r}{\sigma_n^3} e^{-\frac{r^2}{2\sigma_n^2}} : \sum_1^n \frac{k_n}{\sigma_n}.$$

7. In § 3 we denoted by

$$f(r) = \frac{rF(r)}{r_0}$$

the relative frequency function of the true diameters of the corpuscles, represented in a random section. Introducing this function the equation (5) takes the form

$$\phi(x) = x \int_x^R f(r) \frac{dr}{r\sqrt{r^2 - x^2}}.$$

In this form the equation is valid also if the corpuscles are not spherical but of ellipsoidal shape. The different corpuscles need not be of the same form, nor must the axes of the different ellipsoids be parallel, provided that as "diameter" of a plane section out of any corpuscle we take the *geometric mean* of its largest and smallest diameters. The apparent diameter x of a corpuscle should thus be defined as the geometric mean of the largest and smallest diameters in the image obtained in a random section, the true diameter r as the geometric mean of the largest and smallest diameters in the central plane parallel to the section plane.

This rather remarkable fact follows from the following proposition: Take any three-axial ellipsoid and place it with the principal axes pointing in any direction. Two horizontal sections are taken in this ellipsoid, one passing through the centre, the other passing at the distance y from the centre. The areas of the two sections we denote by A_0 and A_y. The distance of the horizontal tangent plane from the centre we denote by h. Then we always have

$$A_y = A_0 \left(1 - \frac{y^2}{h^2}\right).$$

The truth of this proposition follows immediately from the property of an ellipsoid, that the areas of parallel sections are proportional to the areas of similarly situated sections in a sphere.

Now, defining the apparent and the true diameters in the above way, we have

$$\pi r^2 = 4A_0; \quad \pi x^2 = 4A_y,$$

and it is seen that we have the relation

$$x = r\sqrt{1 - \frac{y^2}{h^2}},$$

or

$$\frac{y}{h} = \frac{\sqrt{r^2 - x^2}}{r}.$$

But

$$\left|\frac{dy}{h\,dx}\right| dx = \frac{x\,dx}{r\sqrt{r^2 - x^2}}$$

is the probability that an arbitrary section in a corpuscle of central "diameter" r has an apparent "diameter" x to $x + dx$. Hence this probability is identically the same as in the case of a sphere.

8. It was mentioned above, that the "corpuscle problem" has a parallel in stellar statistics. This parallel is the "globular cluster problem."

The Globular Cluster Problem. A globular cluster, i.e. an agglomeration of stars in spherical layers of equal density around a common centre, is projected on the heavens (or a photographic plate). To express the "true" law of density in space in terms of the observed "apparent" law of density in the projection.

Let $f(r)\,dr$ be the relative number of stars at the distance r to $r+dr$ from the centre of the cluster, and let $\phi(x)\,dx$ be the relative number of stars projected on a plane and falling within the circular ring of radius x to $x+dx$. If $\Delta(r)$ is the density function of the cluster and $\delta(x)$ is the density function of the projection we evidently have

$$N.f(r) = 4\pi r^2\,\Delta(r);\quad N\phi(x) = 2\pi x\delta(x),$$

N being the total number of stars in the cluster. But it is immediately seen that we must have (R being the superior limit to the radius vector)

$$\delta(x) = 2\int_0^{\sqrt{R^2-x^2}} \Delta(\sqrt{x^2+y^2})\,dy,$$

or by a simple transformation

$$\delta(x) = 2\int_x^R \Delta(r)\,\frac{r\,dr}{\sqrt{r^2-x^2}}.$$

Hence we get

$$\phi(x) = x\int_x^R f(r)\,\frac{dr}{r\sqrt{r^2-x^2}},$$

which is the same functional relation as in the corpuscle problem. The two problems are thus—although different in origin and formulation—mathematically identical.

9. The methods of solution expanded in §§ 4 and 5 involve computations and constructions, which will certainly be easily mastered by a trained statistician. I am afraid, however, that the anatomists, who generally have neither the mathematical nor the arithmetical training necessary, will find the methods too complicated for application. Furthermore I suppose that the computations involved will be too unwieldy, when it is required to reduce the observations from numerous organs (the material of Hellman embraces several hundred spleens). To obtain a more ready method of reduction I have proceeded as follows:

The observations give directly the distribution of the apparent diameters in the classes 1·5—2·5; 2·5—3·5; 3·5—4·5 mm., etc. We will denote the frequencies in these classes by $f_2, f_3, f_4, f_5, \ldots f_R$. Although diameters in the classes 0—0·5 and 0·5—1·5 are generally not observed and hence the frequencies in these classes will be unknown, we may, for the sake of completeness, take account of them in the following developments, denoting them by $f_{\frac{1}{2}}$ and f_1. We now have, of course,

$$f_{\frac{1}{2}} = kr_0\int_0^{\frac{1}{2}} \phi(x)\,dx;\quad f_i = kr_0\int_{i-\frac{1}{2}}^{i+\frac{1}{2}} \phi(x)\,dx \quad \ldots\ldots\ldots\ldots\ldots(12).$$

The constant k, which as a factor of proportionality will disappear in the final reductions to relative frequencies, here assumes the value NY, where Y is the total area of the observed sections, and N, as before, the number of corpuscles to the unit volume of tissue.

From equation (5) we now have ($i = 1, 2, \ldots$)

$$f_i = k \int_{i-\frac{1}{2}}^{i+\frac{1}{2}} x\,dx \int_x^R F(r) \frac{dr}{\sqrt{r^2 - x^2}} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (13),$$

or, changing the order of integration and integrating over the variable x,

$$f_i = k \int_{i-\frac{1}{2}}^{i+\frac{1}{2}} dr\, F(r) \sqrt{r^2 - (i-0{\cdot}5)^2} + k \int_{i+\frac{1}{2}}^{R} dr\, F(r) \left(\sqrt{r^2 - (i-0{\cdot}5)^2} - \sqrt{r^2 - (i+0{\cdot}5)^2}\right) \quad \ldots (14).$$

For $f_{\frac{1}{2}}$ we get

$$f_{\frac{1}{2}} = k \int_0^{\frac{1}{2}} dr\, F(r) \,.\, r + k \int_{\frac{1}{2}}^{R} dr\, F(r) \left(r - \sqrt{r^2 - 0{\cdot}25}\right) \quad \ldots\ldots\ldots\ldots (15).$$

As only sporadic corpuscles are found with a diameter above 15 mm. we may put $R = 15{\cdot}5$. Furthermore, we put

$$P_{\frac{1}{2}} = k \int_0^{\frac{1}{2}} dr\, F(r); \quad P_n = \int_{n-\frac{1}{2}}^{n+\frac{1}{2}} dr\, F(r),$$

and it will be seen from the systems (14, 15), that if we put

$$f_{\frac{1}{2}} = P_{\frac{1}{2}} a_{00} + \sum_{j=1}^{15} P_j a_{0j}, \quad f_i = \sum_{j=0}^{15-i} P_{i+j} a_{ij} \quad \ldots\ldots\ldots\ldots\ldots\ldots (16),$$

the coefficients a_{ij} will be approximately given by

$$a_{00} = 0{\cdot}25; \quad a_{0j} = j - \sqrt{j^2 - 0{\cdot}25}; \quad a_{i0} = \sqrt{i^2 - (i-0{\cdot}5)^2};$$
$$a_{ij} = \sqrt{(i+j)^2 - (i-0{\cdot}5)^2} - \sqrt{(i+j)^2 - (i+0{\cdot}5)^2} \quad \ldots\ldots\ldots\ldots (16^{bis}).$$

Now, the system (16) is an approximative system of equations. The nature of the approximations involved will be clear if it is observed that the system has the following "exact" meaning: If determined from the system (16) with the coefficients (16^{bis}) the quantities $P_{\frac{1}{2}}; P_1; P_2; \ldots P_{15}$ are proportional to the numbers of corpuscles having their diameters exactly equal to $\frac{1}{2}; 1; 2; 3; \ldots 15$ mm. respectively, which, if cut by a random plane section, will give the observed distribution $f_{\frac{1}{2}}; f_1; f_2; \ldots f_{15}$ of apparent diameters. Geometrical considerations indeed show that the coefficients a_{ij} are numerically equal to the combined thickness of the two layers parallel to the section plane, in which the centres of corpuscles of diameter j must lie, in order that the apparent diameter be between $i - 0{\cdot}5$ and $i + 0{\cdot}5$.

The use of the system (18) will of course be the following: Inserting the observed values of f_i the corresponding values of P_i are obtained. When, as in the material of Hellman is the case, we do not know $f_{\frac{1}{2}}$ and f_1, it is seen that we can still use all the equations (18) except the first two. Having obtained the values of

Determining the numerical values of the coefficients, the system (16) becomes:

$$
\begin{aligned}
0{\cdot}2500P_1+0{\cdot}1340P_1+0{\cdot}0635P_2+0{\cdot}0420P_3+0{\cdot}0314P_4+0{\cdot}0251P_5+0{\cdot}0209P_6+0{\cdot}0179P_7+0{\cdot}0156P_8+0{\cdot}0139P_9+0{\cdot}0125P_{10}+0{\cdot}0114P_{11}+0{\cdot}0104P_{12}+0{\cdot}0096P_{13}+0{\cdot}0089P_{14}+0{\cdot}0084P_{15}&=f_1\\
0{\cdot}8660P_1+0{\cdot}6136P_2+0{\cdot}3599P_3+0{\cdot}2605P_4+0{\cdot}2052P_5+0{\cdot}1696P_6+0{\cdot}1447P_7+0{\cdot}1263P_8+0{\cdot}1120P_9+0{\cdot}1006P_{10}+0{\cdot}0914P_{11}+0{\cdot}0838P_{12}+0{\cdot}0772P_{13}+0{\cdot}0717P_{14}+0{\cdot}0668P_{15}&=f_1\\
1{\cdot}3229P_2+0{\cdot}9398P_3+0{\cdot}5856P_4+0{\cdot}4396P_5+0{\cdot}3551P_6+0{\cdot}2991P_7+0{\cdot}2588P_8+0{\cdot}2283P_9+0{\cdot}2044P_{10}+0{\cdot}1851P_{11}+0{\cdot}1691P_{12}+0{\cdot}1558P_{13}+0{\cdot}1444P_{14}+0{\cdot}1346P_{15}&=f_2\\
1{\cdot}6583P_3+1{\cdot}1860P_4+0{\cdot}7594P_5+0{\cdot}5810P_6+0{\cdot}4761P_7+0{\cdot}4056P_8+0{\cdot}3542P_9+0{\cdot}3150P_{10}+0{\cdot}2838P_{11}+0{\cdot}2585P_{12}+0{\cdot}2374P_{13}+0{\cdot}2197P_{14}+0{\cdot}2042P_{15}&=f_3\\
1{\cdot}9365P_4+1{\cdot}3913P_5+0{\cdot}9048P_6+0{\cdot}7003P_7+0{\cdot}5793P_8+0{\cdot}4974P_9+0{\cdot}4372P_{10}+0{\cdot}3909P_{11}+0{\cdot}3539P_{12}+0{\cdot}3237P_{13}+0{\cdot}2982P_{14}+0{\cdot}2769P_{15}&=f_4\\
2{\cdot}1794P_5+1{\cdot}5707P_6+1{\cdot}0318P_7+0{\cdot}8049P_8+0{\cdot}6703P_9+0{\cdot}5787P_{10}+0{\cdot}5111P_{11}+0{\cdot}4589P_{12}+0{\cdot}4171P_{13}+0{\cdot}3827P_{14}+0{\cdot}3538P_{15}&=f_5\\
2{\cdot}3979P_6+1{\cdot}7320P_7+1{\cdot}1458P_8+0{\cdot}8990P_9+0{\cdot}7523P_{10}+0{\cdot}6522P_{11}+0{\cdot}5783P_{12}+0{\cdot}5209P_{13}+0{\cdot}4748P_{14}+0{\cdot}4368P_{15}&=f_6\\
2{\cdot}5981P_7+1{\cdot}8798P_8+1{\cdot}2500P_9+0{\cdot}9849P_{10}+0{\cdot}8274P_{11}+0{\cdot}7196P_{12}+0{\cdot}6399P_{13}+0{\cdot}5780P_{14}+0{\cdot}5281P_{15}&=f_7\\
2{\cdot}7839P_8+2{\cdot}0169P_9+1{\cdot}3466P_{10}+1{\cdot}0646P_{11}+0{\cdot}8970P_{12}+0{\cdot}7822P_{13}+0{\cdot}6973P_{14}+0{\cdot}6312P_{15}&=f_8\\
2{\cdot}9580P_9+2{\cdot}1453P_{10}+1{\cdot}4368P_{11}+1{\cdot}1391P_{12}+0{\cdot}9621P_{13}+0{\cdot}8408P_{14}+0{\cdot}7510P_{15}&=f_9\\
3{\cdot}1225P_{10}+2{\cdot}2666P_{11}+1{\cdot}5219P_{12}+1{\cdot}2092P_{13}+1{\cdot}0234P_{14}+0{\cdot}8961P_{15}&=f_{10}\\
3{\cdot}2787P_{11}+2{\cdot}3817P_{12}+1{\cdot}6027P_{13}+1{\cdot}2757P_{14}+1{\cdot}0814P_{15}&=f_{11}\\
3{\cdot}4278P_{12}+2{\cdot}4915P_{13}+1{\cdot}6796P_{14}+1{\cdot}3391P_{15}&=f_{12}\\
3{\cdot}5707P_{13}+2{\cdot}5967P_{14}+1{\cdot}7533P_{15}&=f_{13}\\
3{\cdot}7081P_{14}+2{\cdot}6977P_{15}&=f_{14}\\
3{\cdot}8406P_{15}&=f_{15}
\end{aligned}
\qquad\qquad\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(17).
$$

As the quantities P are here the unknowns and the quantities f may be regarded as known, we must solve this system of linear equations. The solution is found to be:

$$
\begin{aligned}
4{\cdot}0000f_1-0{\cdot}6189f_1+0{\cdot}0951f_2-0{\cdot}0209f_3+0{\cdot}0024f_4-0{\cdot}0012f_5-0{\cdot}0002f_6-0{\cdot}0002f_7-0{\cdot}0001f_8-0{\cdot}0001f_9-0{\cdot}0001f_{10}-0{\cdot}0001f_{11}-0{\cdot}0000f_{12}-0{\cdot}0000f_{13}-0{\cdot}0000f_{14}-0{\cdot}0000f_{15}&=P_1\\
1{\cdot}1547f_1-0{\cdot}5356f_2+0{\cdot}0529f_3-0{\cdot}0258f_4-0{\cdot}0027f_5-0{\cdot}0037f_6-0{\cdot}0019f_7-0{\cdot}0014f_8-0{\cdot}0009f_9-0{\cdot}0007f_{10}-0{\cdot}0005f_{11}-0{\cdot}0004f_{12}-0{\cdot}0003f_{13}-0{\cdot}0003f_{14}-0{\cdot}0002f_{15}&=P_1\\
0{\cdot}7559f_2-0{\cdot}4284f_3+0{\cdot}0338f_4-0{\cdot}0248f_5-0{\cdot}0047f_6-0{\cdot}0047f_7-0{\cdot}0026f_8-0{\cdot}0019f_9-0{\cdot}0014f_{10}-0{\cdot}0010f_{11}-0{\cdot}0007f_{12}-0{\cdot}0006f_{13}-0{\cdot}0005f_{14}-0{\cdot}0004f_{15}&=P_2\\
0{\cdot}6039f_3-0{\cdot}3693f_4+0{\cdot}0256f_5-0{\cdot}0236f_6-0{\cdot}0054f_7-0{\cdot}0051f_8-0{\cdot}0030f_9-0{\cdot}0022f_{10}-0{\cdot}0016f_{11}-0{\cdot}0012f_{12}-0{\cdot}0009f_{13}-0{\cdot}0008f_{14}-0{\cdot}0006f_{15}&=P_3\\
0{\cdot}5164f_4-0{\cdot}3297f_5+0{\cdot}0211f_6-0{\cdot}0223f_7-0{\cdot}0057f_8-0{\cdot}0052f_9-0{\cdot}0032f_{10}-0{\cdot}0024f_{11}-0{\cdot}0017f_{12}-0{\cdot}0014f_{13}-0{\cdot}0010f_{14}-0{\cdot}0009f_{15}&=P_4\\
0{\cdot}4588f_5-0{\cdot}3006f_6+0{\cdot}0181f_7-0{\cdot}0212f_8-0{\cdot}0058f_9-0{\cdot}0052f_{10}-0{\cdot}0033f_{11}-0{\cdot}0025f_{12}-0{\cdot}0018f_{13}-0{\cdot}0014f_{14}-0{\cdot}0011f_{15}&=P_5\\
0{\cdot}4170f_6-0{\cdot}2780f_7+0{\cdot}0161f_8-0{\cdot}0202f_9-0{\cdot}0058f_{10}-0{\cdot}0051f_{11}-0{\cdot}0033f_{12}-0{\cdot}0025f_{13}-0{\cdot}0019f_{14}-0{\cdot}0015f_{15}&=P_6\\
0{\cdot}3849f_7-0{\cdot}2600f_8+0{\cdot}0146f_9-0{\cdot}0193f_{10}-0{\cdot}0058f_{11}-0{\cdot}0050f_{12}-0{\cdot}0033f_{13}-0{\cdot}0025f_{14}-0{\cdot}0019f_{15}&=P_7\\
0{\cdot}3592f_8-0{\cdot}2449f_9+0{\cdot}0134f_{10}-0{\cdot}0185f_{11}-0{\cdot}0057f_{12}-0{\cdot}0050f_{13}-0{\cdot}0033f_{14}-0{\cdot}0025f_{15}&=P_8\\
0{\cdot}3381f_9-0{\cdot}2323f_{10}+0{\cdot}0124f_{11}-0{\cdot}0179f_{12}-0{\cdot}0056f_{13}-0{\cdot}0049f_{14}-0{\cdot}0032f_{15}&=P_9\\
0{\cdot}3203f_{10}-0{\cdot}2214f_{11}+0{\cdot}0116f_{12}-0{\cdot}0172f_{13}-0{\cdot}0054f_{14}-0{\cdot}0048f_{15}&=P_{10}\\
0{\cdot}3050f_{11}-0{\cdot}2119f_{12}+0{\cdot}0110f_{13}-0{\cdot}0166f_{14}-0{\cdot}0053f_{15}&=P_{11}\\
0{\cdot}2917f_{12}-0{\cdot}2036f_{13}+0{\cdot}0104f_{14}-0{\cdot}0161f_{15}&=P_{12}\\
0{\cdot}2801f_{13}-0{\cdot}1961f_{14}+0{\cdot}0099f_{15}&=P_{13}\\
0{\cdot}2697f_{14}-0{\cdot}1894f_{15}&=P_{14}\\
0{\cdot}2604f_{15}&=P_{15}
\end{aligned}
\qquad\qquad\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(18).
$$

$P_2, P_3, \ldots P_{15}$ it will, however, be possible to find the value of $P_{\frac{1}{2}}$ and P_1 by a graphical extrapolation—or rather interpolation, as the curve may be assumed to pass through the origin. Finally, dividing the quantities P by their sum S we have

$$F_{\frac{1}{2}} = \frac{P_{\frac{1}{2}}}{S}, \quad F_i = \frac{P_i}{S},$$

where $F_{\frac{1}{2}}, F_1, F_2, F_3, \ldots F_{15}$ denote the relative frequencies of the true diameters in the classes 0—0·5, 0·5—1·5, 1·5—2·5, 2·5—3·5, ... 14·5—15·5.

In the next section we shall give some numerical examples of the use of equations (18).

10. *Numerical examples.* Of the following numerical examples the first three are arranged in order to obtain a practical test of the mechanical solution (18). In these examples the exact answer is known beforehand, and the examples have been so chosen, that they illustrate the applicability of the method in some typical cases.

Example 1. As a first example I take the following distribution:

$$\phi(x) = \frac{x}{(5)^2} e^{-\frac{x^2}{2(5)^2}}.$$

We here evidently have in *pro mille*

$$f_{\frac{1}{2}} = 1000 \int_0^{\frac{1}{2}} \frac{x}{25} e^{-\frac{x^2}{50}} dx = 1000 \left(1 - e^{-\frac{(0\cdot5)^2}{50}}\right),$$

and

$$f_i = 1000 \int_{i-\frac{1}{2}}^{i+\frac{1}{2}} \frac{x}{25} e^{-\frac{x^2}{50}} dx = 1000 \left(e^{-\frac{(i-\frac{1}{2})^2}{50}} - e^{-\frac{(i+\frac{1}{2})^2}{50}}\right) \quad \ldots\ldots\ldots(19).$$

As was shown in § 6 the true distribution is here

$$F(r) = \frac{r}{5^2} e^{-\frac{r^2}{2\cdot5^2}};$$

and consequently, if the system (18) is applied, the $F_i (= P_i : \Sigma P_i)$ found should be approximately equal to the corresponding f_i.

From (19) we get, writing the numerical values below the symbols,

$f_{\frac{1}{2}}$	f_1	f_2	f_3	f_4	f_5	f_6	f_7	f_8	f_9	f_{10}	f_{11}	f_{12}	f_{13}	f_{14}	f_{15}		Σf
5	39	73	100	116	121	117	105	89	71	54	39	27	18	11	7	8	1000

The residual number (= 8) is of course the value of the integral

$$\int_{15\cdot5}^{\infty} \phi(x)\, dx.$$

It may here be left entirely out of consideration.

Inserting the above values of the f_i in the system (18) we get

$P_{\frac{1}{2}}$	P_1	P_2	P_3	P_4	P_5	P_6	P_7	P_8	P_9	P_{10}	P_{11}	P_{12}	P_{13}	P_{14}	P_{15}
0·8	6·9	11·9	16·3	19·0	19·7	19·0	16·9	14·3	11·3	8·6	6·2	4·2	2·9	1·7	1·8

of which the sum is 161·5. Expressing this in *pro mille* we get

$F_{\frac{1}{2}}$	F_1	F_2	F_3	F_4	F_5	F_6	F_7	F_8	F_9	F_{10}	F_{11}	F_{12}	F_{13}	F_{14}	F_{15}	ΣF
5	43	74	101	118	122	118	105	89	70	53	38	26	18	11	11	1000+2

It is seen that these values of F_i are, indeed, very nearly equal to the f_i, which shows that in the case treated the system (18) has worked very well. That there should be slight deviations at the extremes was, of course, to be expected.

Finally, it follows from § 6 that here we have

$$r_0 = 5 \cdot \sqrt{\frac{\pi}{2}} = 6{\cdot}26.$$

It follows, however, from § 9 that we also have

$$\frac{1}{r_0} = \frac{\Sigma P_i}{\Sigma f_i} = 0{\cdot}1608,$$

which gives in this case

$$r_0 = 6{\cdot}22.$$

Example 2. It is evident that the system (18) can be used also when we have fewer that 15 classes. We then must of course put the superfluous f_i equal to zero. It is also clear, however, that the fewer the classes the less exact will the mechanical solution be. In order to test the degree of approximation in a rather extreme case of 6 classes only, we shall apply the method to the distribution

$$\phi(x) = \frac{x}{(1{\cdot}27)^2} e^{-\frac{x^2}{2(1{\cdot}27)^2}},$$

which gives in *pro mille*

$f_{\frac{1}{2}}$	f_1	f_2	f_3	f_4	f_5	Σf
73	428	355	121	20	2	1000 – 1.

Equations (18) now give

$P_{\frac{1}{2}}$	P_1	P_2	P_3	P_4	P_5	ΣP
57	309	218	64	10	1·0	659

or in *pro mille*

$F_{\frac{1}{2}}$	F_1	F_2	F_3	F_4	F_5	ΣF
87	468	331	98	15	1·5	1000·5.

As an exact solution would, here also, give the apparent distribution back again, we find that 6 classes are generally too few if a higher degree of approximation is required. For the anatomic praxis here in view the accuracy required is, however, not so great but that the application of (18) to series of 6 classes will give fairly satisfactory results.

The exact value of r_0 is here

$$r_0 = 1{\cdot}27\sqrt{\frac{\pi}{2}} = 1{\cdot}57.$$

The approximate solution gives

$$r_0 = \tfrac{1000}{659} = 1{\cdot}52.$$

Example 3. The accuracy of the solution (18) depends not only on the number of classes into which the observations have been divided, but will be affected also by extreme skewness or abnormal excess (i.e. abnormal clustering in the central classes)

in the observed distribution. As an example of a distribution of this kind we may take the following one, derived from the frequency function

$$\phi(x) = 0{\cdot}65\,\frac{x}{(1{\cdot}27)^2}\,e^{-\frac{x^2}{2(1{\cdot}27)^2}} + 0{\cdot}35\,\frac{x}{(4{\cdot}35)^2}\,e^{-\frac{x^2}{2(4{\cdot}35)^2}}.$$

This function has actually been found by me to express the apparent distribution of the stars within the stellar cluster *Messier* 3.

The function gives

$f_{\frac{1}{2}}$	f_1	f_2	f_3	f_4	f_5	f_6	f_7	f_8	f_9	f_{10}	f_{11}	f_{12}	f_{13}	$f\Sigma$
50	295	264	124	59	49	43	36	28	20	13	8	5	3	1000 − 3

and we find from (18)

$P_{\frac{1}{2}}$	P_1	P_2	P_3	P_4	P_5	P_6	P_7	P_8	P_9	P_{10}	P_{11}	P_{12}	P_{13}	ΣP
39·9	204·1	146·8	52·9	14·1	9·5	7·9	6·7	5·3	3·8	2·4	1·3	0·9	0·8	496·4

or in *pro mille*

$F_{\frac{1}{2}}$	F_1	F_2	F_3	F_4	F_5	F_6	F_7	F_8	F_9	F_{10}	F_{11}	F_{12}	F_{13}	ΣF
80	409	296	107	28	19	16	14	11	8	5	3	2	2	1000

Thus

$$r_0 = \frac{1000}{496{\cdot}4} = 2{\cdot}01.$$

The exact distribution is here (§ 6)

$$F_i = 1692 \int_{i-\frac{1}{2}}^{i+\frac{1}{2}} \left(0{\cdot}65\,\frac{r}{(1{\cdot}27)^3}\,e^{-\frac{r^2}{2(1{\cdot}27)^2}} + 0{\cdot}35\,\frac{r}{(4{\cdot}35)^3}\,e^{-\frac{r^2}{2(4{\cdot}35)^2}}\right) dr,$$

which gives

$$r_0 = 1{\cdot}692\,\sqrt{\frac{\pi}{2}} = 2{\cdot}12,$$

and

$F_{\frac{1}{2}}$	F_1	F_2	F_3	F_4	F_5	F_6	F_7	F_8	F_9	F_{10}	F_{11}	F_{12}	F_{13}
64	377	320	123	35	20	16	14	10	7	5	3	2	1

It is seen that the result of equations (18) is here somewhat inexact for the classes where the apparent diameters are most clustered. This approximation would, however, be quite sufficient for anatomical purposes.

Example 4. As a fourth example we take the following distribution, observed by Hellman in 8 sections from a human spleen. (The diameters are given in mm., as measured with 18 × 18 magnification.)

2·5—3·5	3·5—4·5	4·5—5·5	5·5—6·5	6·5—7·5	7·5—8·5	8·5—9·5	9·5—10·5
$f_3 = 146$	$f_4 = 197$	$f_5 = 210$	$f_6 = 184$	$f_7 = 143$	$f_8 = 95$	$f_9 = 57$	$f_{10} = 31$

10·5—11·5	11·5—12·5	12·5—13·5	13·5—14·5	14·5—1·55	Sum
$f_{11} = 15$	$f_{12} = 7$	$f_{13} = 4$	$f_{14} = 2$	$f_{15} = 1$	1092

The frequencies have been slightly smoothed graphically. For the classes 0—0·5 and 0·5—1·5 no observations have been made; for the class 1·5—2·5 the frequency $f_2 = 52$ was found, but this value is evidently too small (it would give a negative value of P_2), probably owing to the fact that the observations are not complete down to 1·5 mm.

Using, consequently, only f_3 to f_{15} we find, inserting in (18),

P_3	P_4	P_5	P_6	P_7	P_8	P_9	P_{10}	P_{11}	P_{12}	P_{13}	P_{14}	P_{15}	ΣP
15·4	32·1	41·5	36·7	30·4	20·4	12·1	6·6	3·1	1·3	0·6	0·4	0·3	200·9

Plotting these values graphically, we may venture an interpolation of the values of $P_{\frac{1}{4}}$, P_1 and P_2, in continuing the curve down to the origin. We have thus found $P_{\frac{1}{4}} = 1$, $P_1 = 3$, and $P_2 = 8$. The sum now becomes 212·9 and we get *pro mille*

$F_{\frac{1}{4}}$	F_1	F_2	F_3	F_4	F_5	F_6	F_7	F_8	F_9	F_{10}	F_{11}	F_{12}	F_{13}	F_{14}	F_{15}
5	14	38	72	151	194	172	143	96	57	31	15	6	3	2	1

This gives

$$r_0 = \tfrac{1}{1000}\left(\sum_{1}^{15} r\,F_r + 0{\cdot}25\,F_{\frac{1}{4}}\right) = 5{\cdot}82$$

for the mean value of the true diameter.

THE FURTHER THEORY OF FRANCIS GALTON'S INDIVIDUAL DIFFERENCE PROBLEM.

By J. O. IRWIN, M.A., M.Sc.

CONTENTS.

	PAGE
Introduction	100
Mean Difference between the pth and qth Individuals in Order of Magnitude of a given Character	102
Moments of the Frequency Distribution of Differences between Consecutive Individuals	104
Actual Frequency Distribution of Differences between consecutive Individuals	108
Approximations to the Frequency Distribution of Differences between Consecutive Individuals	116
Conclusion, including Consideration of the Probable Errors of Correlation Coefficients calculated by use of the Method	120

Introduction.

In August 1902 Sir Francis Galton published a memoir entitled "The Most Suitable Proportion between the Values of First and Second Prizes." He concluded that if the amount of the prizes had to be decided before the competition, it seemed reasonable to divide the prize money in the ratio of the mean excess in merit (supposing it measurable) of candidates who are first over those who are second to the mean excess of those who are second over those who are third.

This led Professor Pearson to consider the following problem:

"A random sample of n individuals is taken from a population of N members, which when N is very large may be taken to obey some law of frequency which is expressed by the curve $y = N\phi(x)$, $y\,\delta x$ being the total frequency of individuals with characters or organs lying between x and $x + \delta x$. It is required to find an expression for the average difference in character between the pth and $(p+1)$th individuals when the sample is arranged in order of magnitude of the character."

This general problem was termed by him "Francis Galton's Individual Difference Problem in Statistics." If $\chi_{n,p}$ be the average difference it is required to find, the solution of the problem is given by

$$\chi_{n,p} = \frac{n!}{(n-p)!\,p!}\int_{-\infty}^{\infty} \alpha^{n-p}(1-\alpha)^p\,dx,$$

where

$$\alpha = \int_{-\infty}^{x} \phi(x)\,dx \qquad \text{(i)}.$$

This result was obtained independently by Professor Pearson and Dr Sheppard and some numerical results are given in the original paper*, on the assumption that the original population is normally distributed when (i) becomes

$$\chi_{n,p} = \frac{\sigma \,.\, n!}{(n-p)!\, p!} \int_{-\infty}^{\infty} \alpha^{n-p} (1-\alpha)^p \, dx,$$

where

$$\alpha = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2} dx \quad \text{.....(ii),}$$

σ being the standard deviation of the original population, and particular values of α being obtainable from the tables of the probability integral.

Two methods were given for the computation:

(α) Quadrature.

(β) An expansion of the integral in a series about the mode of the integrand by a method due to Laplace.

The problem was regarded as a new and important departure in statistical theory, in fact Professor Pearson concluded his investigation in these words: "It (the method) opens up, indeed, many new methods of enquiry, the effectiveness of which however can only be tested by their application in actual statistical practice†."

In order to consider the likelihood of the method being useful in statistics, it is necessary to go rather further into the theory than has hitherto been attempted.

It seems desirable to find not only the mean difference between the pth and $(p+1)$th individuals, but also the standard deviation of that difference, and even the actual frequency distribution of differences, since if it should turn out to be

* *Biometrika*, Vol. I. pp. 385–399.

† It was suggested by Professor Pearson that the above result might be used to obtain fraternal correlation coefficients in the following manner:

Pairs of individuals are drawn at random from the general population and their mean difference for the character in question is measured. Hence from (i) we may find

$$\chi_{2,1} = \lambda\sigma,$$

σ being the standard deviation of the original population, λ a known number calculated from (i) and $\chi_{2,1}$ the mean of the differences between the pairs drawn. If now pairs of individuals be taken not quite at random, but from correlated groups, e.g. pairs of brothers taken at random, the result will be λ times the standard deviation of the correlated groups, e.g. a group of brethren, or we have

$$\chi'_{2,1} = \lambda\sigma\sqrt{1-r^2},$$

where r is the correlation of the group, e.g. the correlation between brothers.

Whence

$$\sqrt{1-r^2} = \frac{\chi'_{2,1}}{\chi_{2,1}}.$$

Now if the phrase "pairs of brothers taken at random" means "brothers" in the ordinary sense, e.g. pairs of individuals who are actually brothers of one another, I do not think this expression for r is correct. It would however be correct on the assumption of linear regression and homoscedasticity, if by the term "brothers" we meant individuals each of whom had a brother with the same value of the character, i.e. if we took pairs of individuals at random, each pair from the same array of the correlation table.

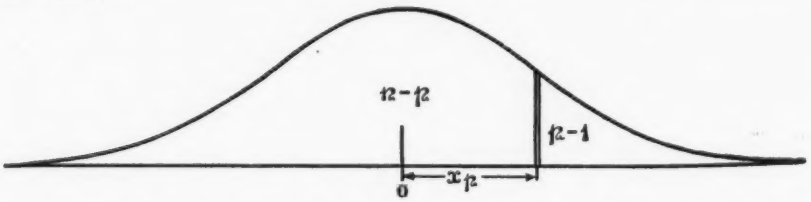
But this is of no practical importance as if we had enough material to form a correlation table we should use the product-moment method, certainly no one would under these circumstances average the differences between pairs taken from the same array.

At the end of this paper, p. 126, we give the correct expression for r when we average differences between pairs of "brothers," in the ordinary sense of the word.

markedly different from the Gaussian that fact would have to be taken into account in any applications of the method. It is the object of this paper to determine that frequency distribution approximately and if possible exactly, on the hypothesis that the original population is normally distributed.

(1) *Mean Difference between the* p*th and* q*th Individuals in Order of Magnitude of a given Character.*

Although not directly concerned with the main point of this investigation, it may be of interest to have a direct formula for the mean difference between the pth and qth individuals in random samples of n, arranged in order of magnitude of the character.



Assuming a frequency distribution $y = N\phi(x)$ in our original population of N (where N is very large), let x_p, x_q be the magnitude of the characters of the pth and qth individuals; then

$$\alpha = \int_{-\infty}^{x_p} \phi(x)\,dx = \text{chance of an individual having a value of the character} < x_p,$$

and $1-\alpha =$ chance of an individual with a greater value of the character.

Therefore chance of the pth individual having his value of the character between x_p and $x_p + dx_p$

$$= \frac{n!}{(n-p)!\,(p-1)!}\,\alpha^{n-p}(1-\alpha)^{p-1}\phi(x_p)\,dx_p,$$

and mean $x_p = \dfrac{n!}{(n-p)!\,(p-1)!}\displaystyle\int_{-\infty}^{\infty} \alpha^{n-p}(1-\alpha)^{p-1}\phi(x_p)\,x_p\,dx_p$

$$= \frac{n!}{(n-p)!\,(p-1)!}\int_{-\infty}^{\infty} \alpha^{n-p}(1-\alpha)^{p-1}\,x_p\,\frac{d\alpha}{dx_p}\,dx_p$$

$$= \frac{n!}{(n-p)!\,(p-1)!}\left\{-\int_0^{\infty}\frac{d}{dx_p}\left(\int_{x_p}^{\infty}(1-\alpha)^{p-1}\alpha^{n-p}\frac{d\alpha}{dx_p}\,dx_p\right)x_p\,dx_p\right.$$
$$\left.+\int_{-\infty}^{0}\frac{d}{dx_p}\left(\int_{-\infty}^{x_p}(1-\alpha)^{p-1}\alpha^{n-p}\frac{d\alpha}{dx_p}\,dx_p\right)x_p\,dx_p\right\}$$

$$= \frac{n!}{(n-p)!\,(p-1)!}\left\{\left[-\int_{x_p}^{\infty}(1-\alpha)^{p-1}\alpha^{n-p}\frac{d\alpha}{dx_p}\,dx_p(x_p)\right]_0^{\infty}\right.$$
$$+\left[\int_{-\infty}^{x_p}(1-\alpha)^{p-1}\alpha^{n-p}\frac{d\alpha}{dx_p}\,dx_p(x_p)\right]_{-\infty}^{0}$$
$$+\int_0^{\infty}dx_p\int_{x_p}^{\infty}(1-\alpha)^{p-1}\alpha^{n-p}\frac{d\alpha}{dx_p}\,dx_p$$
$$\left.-\int_{-\infty}^{0}dx_p\int_{-\infty}^{x_p}(1-\alpha)^{p-1}\alpha^{n-p}\frac{d\alpha}{dx_p}\,dx_p\right\}.$$

All the integrals in this expression are convergent (provided the frequency distribution is of finite range or behaves at $\pm\infty$ like e^{-x^2}) and the terms in square brackets vanish at both limits.

Similarly we have:

$$\text{mean } x_q = \frac{n!}{(n-q)!\,(q-1)!}\left(\int_0^\infty dx_q \int_{x_q}^\infty (1-\alpha)^{q-1}\alpha^{n-q}\frac{d\alpha}{dx_q}dx_q - \int_{-\infty}^0 dx_q\int_{-\infty}^{x_q}(1-\alpha)^{q-1}\alpha^{n-q}\frac{d\alpha}{dx_q}dx_q\right).$$

Therefore mean $(x_p - x_q)$*

$$= \int_0^\infty \left\{\frac{\int_x^\infty (1-\alpha)^{p-1}\alpha^{n-p}\frac{d\alpha}{dx}dx}{B(n-p+1,p)} - \frac{\int_x^\infty (1-\alpha)^{q-1}\alpha^{n-q}\frac{d\alpha}{dx}dx}{B(n-q+1,q)}\right\}dx$$

$$- \int_{-\infty}^0 \left\{\frac{\int_{-\infty}^x (1-\alpha)^{p-1}\alpha^{n-p}\frac{d\alpha}{dx}dx}{B(n-p+1,p)} - \frac{\int_{-\infty}^x (1-\alpha)^{q-1}\alpha^{n-q}\frac{d\alpha}{dx}dx}{B(n-q+1,q)}\right\}dx$$

$$= \int_0^\infty \left\{1 - \frac{\int_{-\infty}^x (1-\alpha)^{p-1}\alpha^{n-p}\frac{d\alpha}{dx}dx}{B(n-p+1,p)} - \left(1 - \frac{\int_{-\infty}^x (1-\alpha)^{q-1}\alpha^{n-q}\frac{d\alpha}{dx}dx}{B(n-q+1,q)}\right)\right\}dx$$

$$- \int_{-\infty}^0 \left\{\frac{\int_{-\infty}^x (1-\alpha)^{p-1}\alpha^{n-p}\frac{d\alpha}{dx}dx}{B(n-p+1,p)} - \frac{\int_{-\infty}^x (1-\alpha)^{q-1}\alpha^{n-q}\frac{d\alpha}{dx}dx}{B(n-q+1,q)}\right\}dx$$

$$= \int_{-\infty}^\infty \left\{\frac{\int_0^a (1-\alpha)^{q-1}\alpha^{n-q}\,d\alpha}{B(n-q+1,q)} - \frac{\int_0^a (1-\alpha)^{p-1}\alpha^{n-p}\,d\alpha}{B(n-p+1,p)}\right\}dx$$

$$= \int_{-\infty}^\infty \{I_a(n-q+1,q) - I_a(n-p+1,p)\}\,dx \quad \ldots\ldots\ldots\ldots\ldots\ldots\text{(iii)},$$

* To prove mean $(x_p - x_q) = \text{mean } x_p - \text{mean } x_q$ we may proceed as follows:

Let δ be the difference between the values of the character in the pth and qth individuals in any *particular sample* of n individuals. Then

$$\delta = x_p - x_q.$$

This is true for every sample, therefore summing for all possible samples and dividing by their number

$$\bar{\delta} = \bar{x}_p - \bar{x}_q.$$

But to some minds it may seem better to approach the problem of this section by the same method as that used for consecutive individuals in the 1902 paper. If we do so we are easily led to the expression

$$\frac{n!}{(n-q)!\,(q-p-1)!\,(p-1)!}\int_{-\infty}^\infty\int_{-\infty}^{x_p} a'^{n-q}(a-a')^{q-p-1}(1-a)^{p-1}(x_p-x_q)\,\phi(x_p)\,\phi(x_q)\,dx_p\,dx_q,$$

for the mean difference, where

$$a = \int_{-\infty}^{x_p}\phi(x)\,dx, \quad a' = \int_{-\infty}^{x_q}\phi(x)\,dx.$$

By expanding $(a-a')^{q-p-1}$ and integrating by parts this may be shown to be equal to

$$\frac{n!}{(n-q)!\,(p-1)!}\int_{-\infty}^\infty \left[\sum_{s=0}^{s=q-p-1}\frac{(-1)^s a^{n-q+s+1}\int_a^1 a^{q-p-s-1}(1-a)^{p-1}da}{(q-p-s-1)!\,s!\,(n-q+s+1)}\right]dx,$$

where

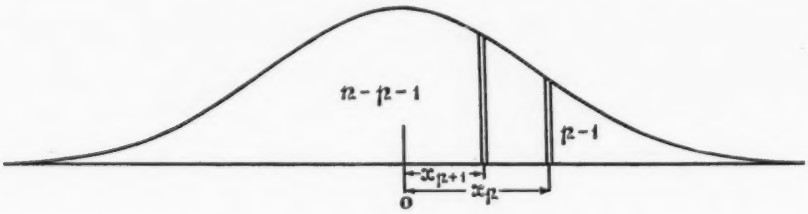
$$a = \int_{-\infty}^x \phi(x)\,dx.$$

where $I_x(r, s)$ denotes the ratio of the incomplete Beta-function $\int_0^x (1-x)^{r-1} x^{s-1}\, dx$ to the complete Beta-function $\int_0^1 (1-x)^{r-1} x^{s-1}\, dx$.

The numerical computation of this expression would involve the calculation of pairs of incomplete Beta-functions of α at convenient intervals of x and then performing a quadrature upon them. It would therefore be very troublesome, and it is questionable whether it would be any improvement on the process of finding the mean differences between the pth, $(p+1)$th, $(p+2)$th, ... qth individuals by formula (i) and summing the results.

(2) *Moments of the Frequency Distribution of Differences between Consecutive Individuals.*

We will now find expressions for the moments about any origin and thence



about the mean of the Frequency Distribution of Differences between consecutive individuals in order of magnitude. A knowledge of the first four moments about the mean will enable us to find an approximate curve which very closely represents the distribution.

Let μ_r' be the rth moment about the origin, and μ_r the rth moment about the mean of the frequency distribution of differences between the pth and $(p+1)$th individuals in random samples of n from the original population $[y = N\phi(x)]$.

Let α denote $\int_{-\infty}^{x_p} \phi(x)\, dx$ and α' denote $\int_{-\infty}^{x_{p+1}} \phi(x)\, dx$.

If we now expand the incomplete Beta-function and perform the summation for every term, the expression may be reduced to

$$\int_{-\infty}^{\infty} \left\{ \frac{n!}{(n-q+1)!\,(q-1)!} (1-\alpha)^{q-1} \alpha^{n-q+1} + \frac{n!}{(n-q+2)!\,(q-2)!} (1-\alpha)^{q-2} \alpha^{n-q+2} + \ldots + \frac{n!\,(1-\alpha)^p \alpha^{n-p}}{p!\,(n-p)!} \right\} dx$$

$$= \text{mean difference between } (q-1)\text{th and } q\text{th individuals}$$
$$+ \text{mean difference between } (q-2)\text{th and } (q-1)\text{th individuals}$$
$$+ \text{mean difference between } (q-3)\text{rd and } (q-2)\text{th individuals}$$
$$+ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$
$$+ \text{mean difference between } (p-1)\text{th and } p\text{th individuals}$$

a result which is easily seen to be equivalent to (iii) on expanding the two incomplete Beta-functions in that expression.

Then

$$\mu_r' = \frac{n!}{(n-p-1)!\,(p-1)!}\int_{-\infty}^{\infty}\int_{-\infty}^{x_p} \alpha'^{n-p-1}(1-\alpha)^{p-1}(x_p - x_{p+1})^r \frac{d\alpha}{dx_p}\frac{d\alpha'}{dx_{p+1}}\,dx_p dx_{p+1}$$

$$= \frac{n!}{(n-p-1)!\,(p-1)!}\int_{-\infty}^{\infty}(1-\alpha)^{p-1}\frac{d\alpha}{dx_p}\,dx_p \left\{\int_{-\infty}^{x_p}\alpha'^{n-p-1}(x_p - x_{p+1})^r \frac{d\alpha'}{dx_{p+1}}\,dx_{p+1}\right\}.$$

Now $$\int_{-\infty}^{x_p}\alpha'^{n-p-1}(x_p - x_{p+1})^r \frac{d\alpha'}{dx_{p+1}}\,dx_{p+1}$$

$$= \left[\frac{\alpha'^{n-p}}{n-p}(x_p - x_{p+1})^r\right]_{-\infty}^{x_p} + r\int_{-\infty}^{x_p}\frac{\alpha'^{n-p}}{n-p}(x_p - x_{p+1})^{r-1}dx_{p+1}.$$

The term in square brackets vanishes at both limits.

Therefore $$\mu_r' = \frac{r\,.\,n!}{(n-p)!\,(p-1)!}\int_{-\infty}^{\infty}(1-\alpha)^{p-1}Q\frac{d\alpha}{dx_p}\,dx_p,$$

where $$Q = \int_{-\infty}^{x_p}\alpha'^{n-p}(x_p - x_{p+1})^{r-1}dx_{p+1}.$$

Therefore $$\mu_r' = \frac{r\,.\,n!}{(n-p)!\,(p-1)!}\left\{\left[-\frac{(1-\alpha)^p}{p}Q\right]_{-\infty}^{\infty} + \int_{-\infty}^{\infty}\frac{(1-\alpha)^p}{p}\frac{dQ}{dx_p}\,dx_p\right\}.$$

The term in square brackets again vanishes at both limits and

$$\frac{dQ}{dx_p} = \int_{-\infty}^{x_p}\alpha'^{n-p}(r-1)(x_p - x_{p+1})^{r-2}dx_{p+1}.$$

Therefore $$\mu_r' = \frac{r(r-1)\,.\,n!}{(n-p)!\,p!}\int_{-\infty}^{\infty}\int_{-\infty}^{x_p}(1-\alpha)^p\,\alpha'^{n-p}(x_p - x_{p+1})^{r-2}dx_p dx_{p+1}$$

$$= \frac{r(r-1)\,.\,n!}{(n-p)!\,p!}\int_{-\infty}^{\infty}\int_{-\infty}^{x}(1-\alpha)^p\,\alpha'^{n-p}(x-y)^{r-2}dx\,dy,$$

where $$\alpha = \int_{-\infty}^{x}\phi(u)\,du, \quad \alpha' = \int_{-\infty}^{y}\phi(u)\,du \qquad \text{.....(iv).}$$

Putting $r = 2, 3, 4$ in succession we have:

$$\left.\begin{aligned}
\mu_2' &= \frac{2\,.\,n!}{(n-p)!\,p!}\int_{-\infty}^{\infty}\int_{-\infty}^{x}(1-\alpha)^p\,\alpha'^{n-p}\,dx\,dy \\
\mu_3' &= \frac{6\,.\,n!}{(n-p)!\,p!}\int_{-\infty}^{\infty}\int_{-\infty}^{x}(1-\alpha)^p\,\alpha'^{n-p}(x-y)\,dx\,dy \\
\mu_4' &= \frac{12\,.\,n!}{(n-p)!\,p!}\int_{-\infty}^{\infty}\int_{-\infty}^{x}(1-\alpha)^p\,\alpha'^{n-p}(x-y)^2\,dx\,dy \\
\text{and of course}\quad \mu_1' &= \frac{n!}{(n-p)!\,p!}\int_{-\infty}^{\infty}\alpha^{n-p}(1-\alpha)^p\,dx
\end{aligned}\right\} \text{.....(v),}$$

and by the ordinary formulae for transferring moments to the mean

$$\left.\begin{aligned}
\mu_2 &= \mu_2' - \mu_1'^2 \\
\mu_3 &= \mu_3' - 3\mu_1'\mu_2' + 2\mu_1'^3 \\
\mu_4 &= \mu_4' - 4\mu_3'\mu_1' + 6\mu_2'\mu_1'^2 - 3\mu_1'^4
\end{aligned}\right\} \text{.....(vi).}$$

We have now to consider the best way of computing the expressions (v) numerically for particular values of r and p on the assumption that the original population is normally distributed.

We have to calculate integrals of the type

$$\int_{-\infty}^{\infty}\int_{-\infty}^{x} Z(x-y)^r\,dxdy,$$

where as many values of Z as desired can be calculated numerically, and r is either 0, 1 or 2.

By turning the axes through 45° this can be transformed into

$$\int_{0}^{\infty}\int_{-\infty}^{\infty} Z'dx'dy',$$

where $$Z' = Z(x-y)^r.$$

The integral is now in a shape in which it is easy to apply a cubature* formula and the process will be aided by the fact that $(x-y)$ is constant when x' is constant. We may note here that throughout all the numerical work in this paper, the standard deviation of the original population is taken as a unit.

The cubature formula we use is the following†:

$$\int_{0}^{\infty}\int_{-\infty}^{\infty} Z\,dxdy = \frac{hk}{120960}[-191u_{-3}+1688u_{-2}-7843u_{-1}+60480u_0$$
$$+128803u_1+119272u_2+121151u_3$$
$$+120960(u_4+u_5+u_6+\ldots)] \quad \ldots\ldots\ldots\ldots\ldots\ldots\text{(vii)},$$

where u_r = sum of all the Z's at intervals k for which $x = rh$. This formula is correct up to and including seventh order differences.

The process, in this case, amounts to calculating ordinates at convenient equal distances, summing them in diagonals and multiplying by the appropriate coefficients. For example, in the case $n = 10$, $p = 1$, values of $(1-\alpha)\alpha'^9$ are calculated at intervals of quarter units of standard deviation from $x = -1{\cdot}75$ to $5{\cdot}00$ and $y = -1{\cdot}00$ to $x + {\cdot}75$ (for smaller values of x and y the ordinates vanish to seven places of decimals).

If we require the second moment of the frequency distribution of differences between the first and second individuals in samples of 10, we have only to sum the successive diagonals beginning from the left to obtain u_{-3}, u_{-2}, u_{-1}, etc.

If we require the third moment the u's are obtained by multiplying the sums of the successive diagonals by $-3, -2, -1, 0, 1, 2, 3$, etc., and dividing each by 4, while for the fourth moment the appropriate multipliers are $\frac{9}{16}, \frac{4}{16}, 0, \frac{1}{16}, \frac{4}{16}, \frac{9}{16}$, etc. The results obtained by the cubature have then to be multiplied by the appropriate factorial terms given in (v) (for the case of $n = 10$, $p = 1$, these will be 20, 60, 120)‡.

* Cubature is the name sometimes given to the process of calculation of volumes which is analogous to quadrature for areas.

† A proof of this formula will be found in *Tracts for Computers*, No. x. p. 27.

‡ The numerical work for this case is given fully in *Tracts for Computers*, No. x. p. 30.

In this manner we have calculated the first four moments of the frequency distribution of differences for the values $n = 3, 10, 50, 100, 500, 1000$ and $p = 1$, $p = 2$ on the assumption that the character is normally distributed in the original population.

We find:

$p = 1$

n	3	10	50	100	500	1000
$\mu_1' = \chi_{n,p}$	·8458	·5388	·3942	·3594	·3044	·2873
μ_2'	1·1730	·5054	·2830	·2378	·1735	·1554
μ_3'	2·1158	·6424	·2825	·2201	·1401	·1196
μ_4'	4·5575	1·0004	·3520	·2558	·1432	·1171

$p = 2$

n	3	10	50	100	500	1000
$\mu_1' = \chi_{n,p}$	·8458	·3453	·2263	·2018	·1655	·1550
μ_2'	1·1730	·2155	·0961	·0772	·0526	·0463
μ_3'	2·1158	·1858	·0581	·0423	·0241	·0201
μ_4'	4·5575	·1990	·0437	·0292	·0141	·0112

while the moments about the mean and β_1, β_2 are as follows:

$p = 1$

n	3	10	50	100	500	1000
$\mu_2 = \sigma_d^2$	·4577	·2151	·1276	·1086	·0808	·0729
μ_3	·3495	·1383	·0704	·0565	·0381	·0331
μ_4	·8988	·2434	·0980	·0736	·0433	·0363
β_1	1·28	1·92	2·39	2·49	2·75	2·83
β_2	4·29	5·26	6·03	6·24	6·63	6·83

$p = 2$

n	3	10	50	100	500	1000
$\mu_2 = \sigma_d^2$	·4577	·0963	·0449	·0365	·0252	·0223
μ_3	·3495	·0449	·0161	·0120	·0071	·0060
μ_4	·8988	·0540	·0127	·0090	·0044	·0037
β_1	1·28	2·26	2·86	2·96	3·15	3·25
β_2	4·29	5·82	6·30	6·76	6·93	7·44

It may also be of interest to give the standard deviations and coefficients of variation $\left(100\frac{\sigma_d}{\mu_1'}\right)$ in the twelve cases.

$p=1$

n	3	10	50	100	500	1000
σ_d	·6765	·4638	·3572	·3295	·2843	·2699
$100\sigma_d/\mu_1'$	79·98	86·08	90·61	91·68	93·40	93·94

$p=2$

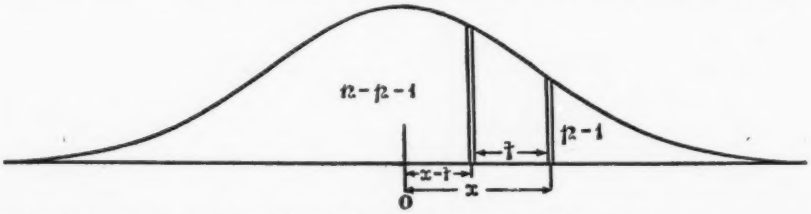
n	3	10	50	100	500	1000
σ_d	·6765	·3103	·2119	·1910	·1588	·1494
$100\sigma_d/\mu_1'$	79·98	89·86	93·64	94·65	95·95	96·39

It will be seen that both for $p=1$ and $p=2$, although the variability decreases fairly rapidly, the coefficient of variation increases steadily and seems to tend to 100 °/₀ as n tends to infinity.

The moments above calculated will be used to obtain curves which approximately represent the frequency distribution of differences.

But before obtaining approximations it will be well to consider more closely their exact nature. It is the purpose of the next section to determine the actual as distinguished from the approximate frequency distributions and where possible to submit them to numerical calculation. Where the latter is not possible our investigation may nevertheless assist us in determining what type of curve may reasonably be used as an approximation.

(3) *Actual Frequency Distribution of Differences between Consecutive Individuals.* The chance of the pth and $(p+1)$th individuals in descending order of magni-



tude (in a sample of n) having a difference between t and $t+dt$ is, on the assumption that the original population is normally distributed,

$$\frac{n!}{(n-p-1)!\,(p-1)!}\int_{x=-\infty}^{x=\infty}\alpha'^{n-p-1}(1-\alpha)^{p-1}\left(\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}(x-t)^2}\right)\left(\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}x^2}\right)dx\,.\,dt$$

$$=\frac{n!}{(n-p-1)!\,(p-1)!}\int_{x=-\infty}^{x=\infty}\alpha'^{n-p-1}(1-\alpha)^{p-1}\frac{1}{2\pi}e^{-\left[\left(x-\frac{1}{2}t\right)^2+\frac{1}{4}t^2\right]}dx\,.\,dt,$$

where $$\alpha' = \int_{-\infty}^{x-t} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}u^2} du, \quad \alpha = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}u^2} du^* \quad \text{.....(viii).}$$

Putting $x - \frac{1}{2}t = x'$ we obtain for the above expression

$$\frac{n!}{(n-p-1)!\,(p-1)!} \int_{x'=-\infty}^{x'=\infty} \alpha'^{n-p-1} (1-\alpha)^{p-1} \frac{1}{2\pi} e^{-x'^2} e^{-\frac{1}{4}t^2} dx'.dt \quad \text{.....(ix).}$$

Now $$\alpha' = \int_{-\infty}^{x'-\frac{1}{2}t} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}u^2} du, \quad \alpha = \int_{-\infty}^{x'+\frac{1}{2}t} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}u^2} du.$$

Therefore $$\frac{\partial \alpha'}{\partial t} = -\frac{1}{2} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(x'-\frac{1}{2}t)^2} = -\frac{1}{2} \frac{\partial \alpha'}{\partial x'},$$

$$\frac{\partial \alpha}{\partial t} = \frac{1}{2} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(x'+\frac{1}{2}t)^2} = \frac{1}{2} \frac{\partial \alpha}{\partial x'}.$$

Now if we expand α, α' in powers of t we obtain:

$$\alpha = \alpha_0 + \left(\frac{\partial \alpha}{\partial t}\right)_0 t + \left(\frac{\partial^2 \alpha}{\partial t^2}\right)_0 \frac{t^2}{2!} + \ldots + \left(\frac{\partial^r \alpha}{\partial t^r}\right)_0 \frac{t^r}{r!} + \ldots,$$

$$\alpha' = \alpha_0' + \left(\frac{\partial \alpha'}{\partial t}\right)_0 t + \left(\frac{\partial^2 \alpha'}{\partial t^2}\right)_0 \frac{t^2}{2!} + \ldots + \left(\frac{\partial^r \alpha'}{\partial t^r}\right)_0 \frac{t^r}{r!} + \ldots,$$

where the suffix 0 denotes that t is to be put zero after the differentiation.

But $$\alpha_0 = \int_{-\infty}^{x'} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}u^2} du,$$

$$\left(\frac{\partial \alpha}{\partial t}\right)_0 = \frac{1}{2} \frac{\partial}{\partial x'} \int_{-\infty}^{x'} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}u^2} du = \frac{1}{2} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x'^2},$$

$$\left(\frac{\partial^r \alpha}{\partial t^r}\right)_0 = \left(\frac{1}{2}\right)^r \frac{d^{r-1}}{dx'^{r-1}} \left(\frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x'^2}\right).$$

Similarly $$\left(\frac{\partial^r \alpha'}{\partial t^r}\right)_0 = \left(-\frac{1}{2}\right)^r \frac{d^{r-1}}{dx'^{r-1}} \left(\frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x'^2}\right).$$

Therefore

$$\alpha = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}u^2} du + \left(\frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x'^2}\right) \frac{t}{2} + \frac{1}{2!} \frac{d}{dx'} \left(\frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x'^2}\right) \left(\frac{t}{2}\right)^2 + \ldots$$
$$+ \frac{1}{r!} \frac{d^{r-1}}{dx'^{r-1}} \left(\frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x'^2}\right) \left(\frac{t}{2}\right)^r + \ldots \quad \text{.....(x),}$$

$$\alpha' = \int_{-\infty}^{x'} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}u^2} du - \left(\frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x'^2}\right) \frac{t}{2} + \frac{1}{2!} \frac{d}{dx'} \left(\frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x'^2}\right) \left(\frac{t}{2}\right)^2 - \ldots$$
$$+ \frac{(-1)^r}{r!} \frac{d^{r-1}}{dx'^{r-1}} \left(\frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x'^2}\right) \left(\frac{t}{2}\right)^r + \ldots \quad \text{.....(x } \textit{bis}\text{).}$$

If now we substitute the expressions (x) and (x *bis*) for α and α' in (ix), re-expand and integrate term by term, we obtain the frequency distribution of differences in the form

$$e^{-\frac{1}{4}t^2} \text{ (Series in Ascending Powers of } t\text{).}$$

* We take as our unit of measurement the standard deviation of the original population.

A mathematical expression for the general term may be found by the use of Burmann's Theorem* for expanding one function in powers of another.

Burmann's Theorem may be written

$$f(z) = f(a) + \sum_{r=1}^{\infty} \frac{B_r [F(z)]^r}{r!},$$

where $$B_r = \left[\frac{d^{r-1}}{dz^{r-1}}\left\{\left(\frac{(z-a)}{F(z)}\right)^r f'(z)\right\}\right]_{z=a},$$

and a is a solution of the equation $F(z) = 0$.

Let $$t = F(\alpha') = \phi(1-\alpha),$$

i.e. the functions $F(\alpha')$ and $\phi(1-\alpha)$ are the inverse functions of

$$\alpha' = \int_{-\infty}^{x'-\frac{1}{2}t} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}u^2} du \quad \text{and} \quad 1-\alpha = \int_{x'+\frac{1}{2}t}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}u^2} du\dagger.$$

Then a is given by $$F(\alpha') = \phi(\alpha) = 0,$$

or $$a = \int_{-\infty}^{x'} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}u^2} du.$$

Accordingly $$\alpha'^{n-p-1} = a^{n-p-1} + \sum_{r=1}^{\infty} \frac{B_r t^r}{r!},$$

where $$B_r = \left[\frac{d^{r-1}}{d\alpha'^{r-1}}\left\{\left(\frac{\alpha'-a}{t}\right)^r (n-p-1)\alpha'^{n-p-2}\right\}\right]_{\substack{\alpha'=a,\\ \text{i.e. } t=0}} \quad \text{.....(xi).}$$

Similarly $$(1-\alpha)^{p-1} = (1-a)^{p-1} + \sum_{r=1}^{\infty} \frac{C_r t^r}{r!},$$

where $$C_r = -\left[\frac{d^{r-1}}{d\alpha^{r-1}}\left\{\left(\frac{\alpha-a}{t}\right)^r (p-1)(1-\alpha)^{p-2}\right\}\right]_{\substack{a=a,\\ \text{i.e. } t=0}} \quad \text{.....(xi \textit{bis}).}$$

Therefore $$\alpha'^{n-p-1}(1-\alpha)^{p-1} = U_0 + \sum_{r=1}^{\infty} \frac{U_r t^r}{r!},$$

where $$U_r = B_0 C_r + r B_1 C_{r-1} + \frac{r(r-1)}{2!} B_2 C_{r-2} + \ldots + B_r C_0 \quad \text{.....(xii).}$$

Therefore the frequency distribution is

$$y = \frac{n!}{(n-p-1)!\,(p-1)!} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}t^2} \sum_{0}^{\infty} \frac{V_r t^r}{r!} \quad \text{.....(xiii)}$$

where $$V_r = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} U_r e^{-x'^2} dx'$$

and U_r, B_r, C_r have the above values; also

$$B_0 = a^{n-p-1} = \left\{\int_{-\infty}^{x'} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}u^2} du\right\}^{n-p-1},$$

$$C_0 = (1-a)^{p-1} = \left\{\int_{x'}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}u^2} du\right\}^{p-1},$$

$$U_0 = B_0 C_0.$$

* See Edwards' *Differential Calculus*, 1912 edition, p. 455 and Whittaker and Watson's *Modern Analysis*, p. 129.

† These inverse functions certainly exist, since for a given x', there is only one value of a corresponding to one of t and vice versa.

We thus have a formal solution to the problem; but even if it could be shown theoretically that the series was always convergent, it would still have to be admitted that the form is not easily adaptable to numerical calculation.

The calculation of the B's and C's in the first place involves the evaluation of a number of undetermined forms, which becomes somewhat troublesome when r is greater than 3. In practice, it is easier to obtain as many terms as desired in the expansions of α'^{n-p-1} and $(1-\alpha)^{p-1}$ direct from the expansions for α' and $1-\alpha$, by Arbogast's rule or some other convenient process*. We have, however, succeeded in working out completely the cases of $n=2$ and $n=3$.

(α) $n=2$, $p=1$. In this case $B_0=1$, $C_0=1$, $U_0=1$, $B_r=U_r=C_r=0$ for all positive integral values of r,

$$V_r=\int_{-\infty}^{\infty}\frac{1}{\sqrt{2\pi}}e^{-x'^2}dx'=\frac{1}{\sqrt{2}},$$

and the distribution is

$$y=\sqrt{2}\,\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}t^2}=\frac{1}{\sqrt{\pi}}e^{-\frac{1}{2}t^2} \quad \text{.....(xiv)†.}$$

The range is from 0 to ∞, since the first individual has, by hypothesis, a greater value of the character than the second, and so the differences will always be of the same sign which we take to be positive.

Thus we have the interesting result that the frequency distribution of differences between pairs of individuals taken at random from a normal population, is itself a normal curve whose standard deviation is $\sqrt{2}$ times the standard deviation of the original population.

(β) $n=3, p=1$. Here $B_0=a=\int_{-\infty}^{x'}\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}u^2}du$, $C_0=1$,

$$\alpha'^{n-p-1}=\alpha', \quad (1-\alpha)^{p-1}=1.$$

Therefore $$U_r=B_r=\frac{1}{2^r}(-1)^r\frac{d^{r-1}}{dx'^{r-1}}\left(\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}x'^2}\right), \quad C_r=0,$$

for all positive integral values of r.

* Arbogast's rule in its most convenient form is as follows: if

$$\left(a_0+a_1x+a_2\frac{x^2}{2!}+a_3\frac{x^3}{3!}+\ldots\right)^n$$

be expanded in the form

$$A_0+A_1x+A_2\frac{x^2}{2!}+A_3\frac{x^3}{3!}+\ldots$$

the coefficients $A_1, A_2, A_3, \ldots$ can be successively determined from

$$A_{r+1}=\left(a_1\frac{\partial}{\partial a_0}+a_2\frac{\partial}{\partial a_1}+a_3\frac{\partial}{\partial a_2}+\ldots\right)A_r.$$

See Edwards' *Differential Calculus*, p. 450.

† The moments of this distribution are easily found and are as follows:

$$\mu_1'=\frac{2}{\sqrt{\pi}}, \quad \mu_2'=2, \quad \mu_3'=\frac{8}{\sqrt{\pi}}, \quad \mu_4'=12, \quad \mu_2=2\left(1-\frac{2}{\pi}\right), \quad \mu_3=\frac{4}{\sqrt{\pi}}\left(\frac{4}{\pi}-1\right), \quad \mu_4=4\left(3-\frac{4}{\pi}-\frac{12}{\pi^2}\right).$$

This value of μ_1' is in agreement numerically with that found by Pearson in the 1902 paper.

Whence $$V_r = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} B_r e^{-x'^2} dx'$$

and $$y = \frac{6}{\sqrt{2\pi}} e^{-\frac{1}{2}t^2} \sum_0^\infty \frac{V_r t^r}{r!} \quad \text{.....(xv).}$$

The numerical calculation was carried out as follows:

Writing $$z = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2},$$

we have $$\frac{V_r}{\sqrt{2\pi}} = \int_{-\infty}^{\infty} B_r z^2 dx' \quad \text{.....(xvi),}$$

$$B_r = \frac{1}{2^r}(-1)^r z \left[(-\tfrac{1}{2})^{r-1} (2x')^{r-1} + \frac{(r-1)(r-2)}{1!} (-\tfrac{1}{2})^{r-2} (2x')^{r-2} \right.$$
$$\left. + \frac{(r-1)(r-2)(r-3)(r-4)}{2!} (-\tfrac{1}{2})^{r-3} (2x')^{r-5} + \dots \right] \quad \text{.....(xvii).}$$

From (xvii) the explicit values of the B's were obtained from B_1 to B_{17} and then $\frac{V_r}{\sqrt{2\pi}}$ was calculated by a quadrature, taking ordinates for x' at every half unit of standard deviation. The work was simplified by the fact that when r is even B_r contains only odd powers of x' and consequently $V_r = 0$, that is to say only odd powers of t occur in the series in (xv). It would take up too much space to give here all the numerical work in full, but as a specimen the calculation of $\frac{V_5}{\sqrt{2\pi}}$ is given.

We have $$B_5 = -\tfrac{1}{3\cdot 2} z (x'^4 - 6x'^2 + 3),$$

and $$\frac{V_5}{\sqrt{2\pi}} = -\int_{-\infty}^{\infty} \tfrac{1}{3\cdot 2} z^3 (x'^4 - 6x'^2 + 3)\, dx'.$$

z can be obtained from tables of ordinates of the normal curve, whence we have the following values of x' and $z^3(x'^4 - 6x'^2 + 3)$:

x'	$z^3(x'^4 - 6x'^2 + 3)$
± 4·5	·0000000
± 4·0	·0000000
± 3·5	·0000000
± 3·0	·0000030
± 2·5	·0000246
± 2·0	− ·0007870
± 1·5	− ·0118135
± 1·0	− ·0283346
± ·5	·0681852
0	·1904811

Using the Euler-Maclaurin formula which, in this case, reduces to half the sum of the ordinates, we have:

$$\int_{-\infty}^{\infty} \tfrac{1}{3\cdot 2} z^3 (x'^4 - 6x'^2 + 3)\, dx' = \cdot 1215182$$

and $$\frac{V_5}{\sqrt{2\pi}} = -\cdot 0037974,$$

$$\frac{V_5}{\sqrt{2\pi}\,5!} = -\cdot 0000316.$$

In a similar way the other coefficients of the series in (xv) can be calculated and we find for the frequency distribution

$$y = e^{-\frac{1}{2}t^2}[\cdot 84628 - \cdot 27566t + \cdot 00766t^3 - \cdot 00019t^5 + \cdot 0000036t^7 - \cdot 000000026t^9 + \ldots] \quad \ldots\ldots\text{(xviii)}.$$

This is the actual frequency distribution of the differences between the first and second individuals in samples of three, taken at random from a normal population; the unit in terms of which t is measured being the standard deviation of the original population.

By repeated integration of this series term by term we may obtain series for the moments which are however less rapidly convergent than (xviii). For this purpose it was necessary to calculate all the odd V's up to V_{17}, so that the first four moments might be obtained and compared with the results of Section (2), p. 105.

If we write (xviii)

$$y = e^{-\frac{1}{2}t^2}[a_0 + a_1 t + a_3 t^3 + a_5 t^5 + \ldots]$$

the series for the rth moment is given by

$$\mu_r' = [b_0 + b_1 + b_3 + b_5 + \ldots + b_{2s-1} + \ldots],$$

where

$$b_{2s-1} = a_{2s-1}\int_0^\infty e^{-\frac{1}{2}t^2} t^{r+2s-1}\,dt \quad \ldots\ldots\ldots\ldots\ldots\ldots\text{(xix)}$$

We have the following numerical values:

n	$a_n\int_0^\infty t^n e^{-\frac{1}{2}t^2}dt$	$a_{n-1}\int_0^\infty t^n e^{-\frac{1}{2}t^2}dt$	$a_{n-2}\int_0^\infty t^n e^{-\frac{1}{2}t^2}dt$	$a_{n-3}\int_0^\infty t^n e^{-\frac{1}{2}t^2}dt$	$a_{n-4}\int_0^\infty t^n e^{-\frac{1}{2}t^2}dt$
0	1·49999				
1	−·55132	1·69257			
2	0	−·97720	3·00000		
3	·06128	0	−2·20531	6·77027	
4	0	·16286	0	−5·86322	17·99999
5	−·01216	0	·49006	0	−17·64250
6	0	−·04033	0	1·62864	0
7	·00276	0	−·14569	0	5·88073
8	0	·01072	0	−·54658	0
9	−·00032	0	·04424	0	−2·33312
10	0	−·00138	0	·19296	0
11	−·00014	0	−·00633	0	·93348
12	0	−·00068	0	−·03039	0
13	−·00003	0	−·00339	0	−·15203
14	0	−·00015	0	−·01761	0
15	·00002	0	−·00081	0	−·09487
16	0	·00009	0	−·00452	0
17	−·00001	0	·00054	0	−·02594
18			0	·00318	0
19			−·00020	0	·01937
20				−·00128	0
21					−·00819

whence we have:

$$\text{area} = 1{\cdot}0001, \qquad \mu_3' = 2{\cdot}1315,$$
$$\mu_1' = \cdot 8465, \qquad \mu_4' = 4{\cdot}5769.$$
$$\mu_2' = 1{\cdot}1731,$$

By the method of Section (2) we had:

$$\text{area} = 1{\cdot}0000, \qquad \mu_3' = 2{\cdot}1158,$$
$$\mu_1' = \cdot 8458, \qquad \mu_4' = 4{\cdot}5575.$$
$$\mu_2' = 1{\cdot}1730,$$

The agreement seems satisfactory, as the divergence in μ_3' and μ_4' is due to stopping at the coefficient of t^{17} and neglecting higher powers in our original series (xviii). The moments about the mean by the two methods are given below:

	μ_2	μ_3	μ_4
Method of Section 2	·4565	·3656	·8629
Actual Series	·4577	·3495	·8988

The divergences in μ_3 and μ_4 occur for the same reason as in μ_3' and μ_4'.

We have further calculated from (xviii) the ordinates of the curve at intervals of 0·5, so that it may be plotted.

t	y	t	y
0·0	·84628	3·0	·01958
0·5	·66642	3·5	·00613
1·0	·45022	4·0	·00168
1·5	·26055	4·5	·00044
2·0	·12898	5·0	·00011
2·5	·05458	5·5	·00003
		6·0	·00001

(γ) $n = 10$, $p = 1$. An attempt, extending over some weeks, was made to compute the actual frequency distribution for $n = 10$, $p = 1$.

In this case it may be seen from (ix) and (xiii) that the frequency distribution is:

$$y = \frac{90}{\sqrt{2\pi}} e^{-\frac{1}{2}t^2} \sum_0^\infty \frac{V_r t^r}{r!},$$

where

$$V_r = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} B_r e^{-x'^2} dx' \quad \text{.....(xx)},$$

and $\dfrac{B_r}{r!}$ is the coefficient of t^r in the expansion of α'^s in ascending powers of t, α' being given by (x *bis*). The first ten terms in this expansion were worked out by Arbogast's theorem and the first eight V's were actually computed; but the resulting series was not sufficiently convergent to make the numerical computation of the frequency distribution possible from a knowledge of these eight terms. The attempt was then abandoned owing to the amount of arithmetic involved in the computation of the higher V's becoming prohibitive. To show that this is so we need only mention that B_8 is given by

$$\begin{aligned} B_8 = \tfrac{1}{3^{\frac{1}{2}}} [&5040z^8 + 141120az^6z_1 + 2520a^2(56z^5z_3 + 210z^4z_1^2) \\ &+ 4200a^3(14z^4z_3 + 112z^3z_1z_2 + 84z^2z_1^3) \\ &+ 210a^4(56z^3z_4 + 420z^2z_1z_3 + 280z^2z_2^2 + 840zz_1^2z_2 + 105z_1^4) \\ &+ 42a^5(28z^2z_5 + 168zz_1z_4 + 280zz_2z_3 + 280z_1z_2^2 + 210z_1^2z_3) \\ &+ 7a^6(8zz_6 + 28z_1z_5 + 56z_2z_4 + 35z_3^2) + a^7z_7] \quad \text{.....(xxi)}, \end{aligned}$$

where $a = \int_{-\infty}^{x'} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}u^2} du, \quad z = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x'^2},$

and $z_r = \frac{d^r}{dx'^r}\left(\frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x'^2}\right).$

Sufficient numerical values of B_s have to be calculated in order to make it possible to obtain V_s by a quadrature; further the number of terms in B_r increases with r.

It would thus seem that for the higher values of n we are thrown back on the method of fitting curves to the moments calculated in Section (2) in order to obtain information about the frequency distributions.

This information may however be supplemented in one respect, by the computation of the values of y in the actual curves at the start of their range, i.e. when $t = 0$.

We then have

$$y = \frac{n!}{(n-p-1)!\,(p-1)!} \int_{x'=-\infty}^{x'=\infty} a^{n-p-1}(1-a)^{p-1} z^2 dx',$$

where $a = \int_{-\infty}^{x'} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}u^2} du, \quad z = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x'^2}$ .....(xxii).

For $p = 1$, $n = 10, 50, 100, 500, 1000$ respectively, this reduces to

$$90 \int_{-\infty}^{\infty} a^8 z^2 dx',$$

$$2450 \int_{-\infty}^{\infty} a^{48} z^2 dx',$$

$$9900 \int_{-\infty}^{\infty} a^{98} z^2 dx',$$

$$249{,}500 \int_{-\infty}^{\infty} a^{498} z^2 dx',$$

$$999{,}000 \int_{-\infty}^{\infty} a^{998} z^2 dx',$$

and for $p = 2$, $n = 10, 50, 100, 500, 1000$ to

$$720 \int_{-\infty}^{\infty} a^7 (1-a)\, z^2 dx',$$

$$117{,}600 \int_{-\infty}^{\infty} a^{47} (1-a)\, z^2 dx',$$

$$970{,}200 \int_{-\infty}^{\infty} a^{97} (1-a)\, z^2 dx',$$

$$12{,}425{,}100 \int_{-\infty}^{\infty} a^{497} (1-a)\, z^2 dx',$$

$$997{,}002{,}000 \int_{-\infty}^{\infty} a^{997} (1-a)\, z^2 dx'.$$

These integrals may be calculated by means of a quadrature. This has been done, taking ordinates at intervals of ·2, it being necessary for $n = 100, 500$ and

1000 to take "a" to ten places and use ten figure logarithms. We have the following results:

Values of Ordinate when $t = 0$.

	$n=10$	$n=50$	$n=100$	$n=500$	$n=1000$
$p=1$	1·539	2·249	2·495	3·037	3·241
$p=2$	2·537	4·104	4·656	5·769	6·193

(4) *Approximations to the Frequency Distribution of Differences between Consecutive Individuals.*

We will now turn to the problem of finding approximations to the frequency distributions by fitting curves to the moments calculated in Section (2). Much difficulty has been experienced in finding a curve of a comparatively simple type which is at the same time an adequate representation of the distribution for all the cases that have been considered. There are of course only two cases in which the adequacy of the approximation may be directly tested, namely, $n = 2$ and $n = 3$, for these are the only two cases in which the ordinates may be calculated. In other cases we must be content to compare the first four moments about the mean and the initial ordinates of the actual and approximate curves. This is, however, adequate, as if these five quantities are in good agreement, the fit may be regarded as good.

Three types of curve suggest themselves as possibilities for approximation.

(α) A Pearson frequency-curve.

(β) A curve of the form $e^{-\frac{1}{2}x^2}$ (polynomial in x), i.e. the sum of a *finite number* of tetrachoric functions.

(γ) The tail of a normal curve, i.e. a curve of the form

$$y = y_0 e^{-\frac{1}{2}\left\{\frac{(x+h)^2 - h^2}{\Sigma^2}\right\}}.$$

(α) Turning to the values of β_1, β_2 given on page 107, we see that they lead to the following types:

$p = 1$, $n = 3, 10, 50, 100, 500, 1000$; $p = 2$, $n = 3, 10$ — Type I_1, i.e. $y = y_0\left(1 + \frac{x}{a_1}\right)^{m_1}\left(1 - \frac{x}{a_2}\right)^{m_2}$, where the origin is at the mode and a_1, a_2, m_1, m_2 are positive.

$p = 2$, $n = 50, 100, 500, 1000$ — Type J_1, i.e. $y = y_0\left(1 + \frac{x}{a_1}\right)^{m_1} \Big/ \left(\frac{x}{a_2} - 1\right)^{m_2}$, where the origin is at $x = 0$ and a_1, a_2, m_1, m_2 are positive.

Now these curves start in the one case with a zero ordinate, and in the other case with an infinite ordinate and so are not adapted to the representation of the actual distributions which we know start with a finite ordinate. If their constants be worked out from the moments, the curves may be shown to fit well except near the beginning of the range, but their failure there is a grave disadvantage.

It may be noticed if we examine the β_1, β_2 plane that in all the cases considered the point β_1, β_2 lies near that portion of the biquadratic curve along which the appropriate type is Pearson's Type IX, i.e. $y = y_0\left(1 - \frac{x}{a}\right)^m$. This is a J curve starting with a finite ordinate. It has accordingly been tried for the case of $n = 3$, a and m being determined from the known values of y_0 and μ_1', but the fit while naturally good at the beginning of the range is on the whole worse than before.

This method has therefore to be abandoned.

(β) The next method which suggests itself is to take a curve of the form $e^{-\frac{1}{2}x^2}(a + bx + cx^2 + dx^3)$ or of the form $e^{-\frac{1}{2}x^2}(a + bx + cx^2 + dx^3 + ex^4 + fx^5)$ and to determine the constants from the known values of the initial ordinate, the area and in one case the first two and in the other case the first four moments about $x = 0$.

Now for $n = 3$ both these forms lead to an excellent fit but when the method is tried for higher values of n it gives rise to negative frequencies for large values of x; showing that 3 or 5 terms are not enough adequately to represent what is really an infinite series, even when the coefficients of the terms retained are modified so as to make the moments agree with their actual values. So that this method too is in general inapplicable.

(γ) The fitting of a curve of the form

$$y = y_0 e^{-\frac{1}{2}\left\{\frac{(x+h)^2 - h^2}{\Sigma^2}\right\}}$$

is perhaps suggested by the fact that when $n = 2$ the frequency distribution of differences is accurately of this type and further that it would seem to be possible to expand y in the form $e^{-\frac{1}{2}x^2}$ (Series in powers of x), a form which has actually been found to represent the distribution.

This curve is fitted by making its mean and standard deviation identical with the actual mean and standard deviation, which leads to equations from which y_0, h and Σ may be determined. A comparison between the actual values of y_0, μ_3 and μ_4 and those given by the approximate curve is then a test of the adequacy of the fit in those cases when the ordinates of the actual distribution cannot be directly calculated.

It has been shown by Pearson and Lee* that if $h' = \frac{h}{\Sigma}$, and if $m_r(h')$, or for brevity m_r', represents the incomplete normal moment function $\int_{h'}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2} x^r dx$,

* See *Biometrika*, Vol. VI. p. 63.

and that if $y_0 = \frac{M}{\Sigma}\left(\frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2}\right)$, so that M is the area of the complete normal curve of which our curve is a portion, then

$$\left(\frac{\sigma_d}{\mu_1'}\right)^2 = \psi_1 = \frac{\mu_2' - \mu_1'^2}{\mu_1'^2} = \frac{m_2' m_0' - m_1'^2}{(m_1' - h' m_0')^2},$$

$$M = 1/m_0'^*,$$

$$\Sigma = \mu_1' m_0'/(m_1' - h' m_0') = \mu_1' \psi_2 \text{.....(xxiii).}$$

ψ_1 and ψ_2 have been tabled against h'. Hence we may find ψ_1 from the actual mean and standard deviation and given ψ_1 we may find h' from the table. Then from h' ψ_2 is found and accordingly:

$$\Sigma = \mu_1' \psi_2,$$

and finally

$$y_0 = \frac{M}{\Sigma} Z_{h'} \text{ .....(xxiv).}$$

To test the adequacy of the fit μ_3 and μ_4 can then be calculated from the relations

$$\mu_3 = \Sigma^3 \left\{\frac{m_3'}{m_0'} - 3\frac{m_1'}{m_0'}\cdot\frac{m_2'}{m_0'} + 2\left(\frac{m_1'}{m_0'}\right)^3\right\},$$

$$\mu_4 = \Sigma^4 \left\{\frac{m_4'}{m_0'} - 4\frac{m_3'}{m_0'}\cdot\frac{m_1'}{m_0'} + 6\frac{m_2'}{m_0'}\cdot\left(\frac{m_1'}{m_0'}\right)^2 - 3\left(\frac{m_1'}{m_0'}\right)^4\right\} \text{ .....(xxv),}$$

and together with y_0 compared with their actual values.

If

$$y = y_0 e^{-\frac{1}{2}\left\{\frac{(x+h)^2 - h^2}{\Sigma^2}\right\}} \text{ .....(xxvi)}$$

be fitted in this manner by using the means and standard deviations given on p. 107, we find the following values for y_0, h and Σ:

$p=1$				$p=2$			
n	h	Σ	y_0	n	h	Σ	y_0
2	0	1·414	·564	*2*	—	—	—
3	·525	1·271	·848	*3*	·525	1·271	·848
10	1·295	1·095	1·529	*10*	1·609	·878	2·540
50	2·133	1·060	2·250	*50*	2·295	·785	4·105
100	2·384	1·046	2·508	*100*	2·587	·774	4·655
500	2·896	1·027	3·037	*500*	3·071	·749	5·761
1000	3·116	1·025	3·237	*1000*	3·170	·733	6·148
				1200	3·189	·729	6·289

These values if substituted in (xxvi) give us the approximate frequency distributions of differences between consecutive individuals.

To test the adequacy of the approximations we have the following table comparing the actual values of y_0, μ_3 and μ_4 with the values obtained from (xxiv) and (xxv).

* Here we are supposing the area of the tail to be unity so that the total frequency of our approximate distribution will be unity as it was taken to be in the actual distribution of Section (3). If M' be the area of the tail we have $M = M'/m_0'$.

$p = 1$

	$n=3$		$n=10$		$n=50$		$n=100$		$n=500$		$n=1000$	
	Actual	Fitted	Actual	Fitted	Actual	Fitted	Actual	Fitted	Actual	Fitted	Actual	Fitted
y_0	·848	·848	1·539	1·529	2·249	2·250	2·495	2·508	3·037	3·037	3·241	3·237
μ_3	·350	·352	·138	·136	·070	·069	·057	·057	·038	·036	·033	·034
μ_4	·899	·906	·243	·239	·098	·104	·074	·070	·043	—*	·036	—

$p = 2$

	$n=3$		$n=10$		$n=50$		$n=100$		$n=500$		$n=1000$	
	Actual	Fitted	Actual	Fitted	Actual	Fitted	Actual	Fitted	Actual	Fitted	Actual	Fitted
y_0	·848	·848	2·537	2·540	4·104	4·105	4·656	4·660	5·769	5·761	6·193	6·148
μ_3	·350	·352	·045	·045	·016	·015	·012	·008	·007	·007	·006	·006
μ_4	·899	·906	·054	·055	·013	·017	·009	—	·004	—	·004	—

This shows a satisfactory agreement.

In the case of $n = 3$, $p = 1$ or 2 we may proceed further and compare the ordinates of the actual frequency distribution (xviii) with those of its approximation

$$y = ·848\, e^{-\frac{1}{2}\left\{\frac{(x+·525)^2 - (·525)^2}{(1·271)^2}\right\}}.$$

We find:

x	y	
	Actual	Fitted
0	·846	·848
·5	·666	·667
1·0	·450	·449
1·5	·261	·260
2·0	·129	·128
2·5	·055	·054
3·0	·020	·020
3·5	·006	·006
4·0	·002	·002
4·5	·000	·000

which also shows a satisfactory agreement.

* The values of μ_4 in the approximate curves (when $n=500$ and 1000, $p=1$ and $n=100$, 500 and 1000, $p=2$) cannot be obtained because three decimal places in h are inadequate for this purpose and for such large values of h we cannot table ψ_1 to more than three figures even using Sheppard's Tables of the Probability Integral to twelve figures. Therefore h itself cannot be obtained to more figures.

The approximate distributions are shown in Diagrams 1 and 2, while in Diagrams 3, 4, 5 h, Σ and y_0 are plotted against n. The points lie on smooth curves and thus it is possible to obtain from these graphs the approximate frequency distributions of differences between consecutive individuals for other values of n when p is either 1 or 2; that is the approximate frequency distribution of differences between the first and second or second and third individuals in samples of any size.

In Diagram 1 the unit of the horizontal scale is the standard deviation of the original population, the total frequency is taken as unity, and where the reading of the vertical scale is y the corresponding frequency is $y\delta x$. The standard deviations in Diagram 3 and the h's in Diagrams 4 and 5 are also measured in terms of the standard deviation of the original population as a unit.

Conclusion.

(1) We have shown that the frequency distributions of differences between the first and second and between the second and third individuals in order of magnitude, in samples of any size taken from a normal population, are closely represented by J curves whose equations may be written in the form

$$y = y_0 e^{-\frac{1}{2}\left\{\frac{(x+h)^2 - h^2}{\Sigma^2}\right\}}.$$

We will now consider briefly what is the bearing of this on Galton's prize problem. Galton said that the proper ratio in which to divide the prize was the ratio of the mean excess in merit of the first individual over the second to that of the second over the third.

But it may be urged that we should take not the means but the most probable values of these quantities and find their ratio. This would lead us to a difficulty, for each of these is zero*, or we have to evaluate in some way an undetermined form, and it is not clear how this should be done. For our own part, we should be disposed to keep to the means though fully realising that this is somewhat arbitrary†.

(2) We will now consider how to obtain the correlation between brothers by averaging

(α) The differences, all taken positively, of pairs of individuals drawn at random from the population.

(β) The differences, all taken positively, of pairs of brothers drawn at random.

We shall assume that the original population is normally distributed and that the bivariate correlation surface for pairs of brothers is also normal.

* The idea that the most probable value should occur when the two individuals are exactly alike ceases to be strange when it is realised that the probability of any *particular* value is infinitesimal.

† [The writer we think overlooks what Galton was aiming at; he was not seeking the most probable individual result, but how to be just *on the average* to competitors. Ed.]

APPROXIMATE FREQUENCY DISTRIBUTIONS
OF DIFFERENCES BETWEEN FIRST & SECOND INDIVIDUALS

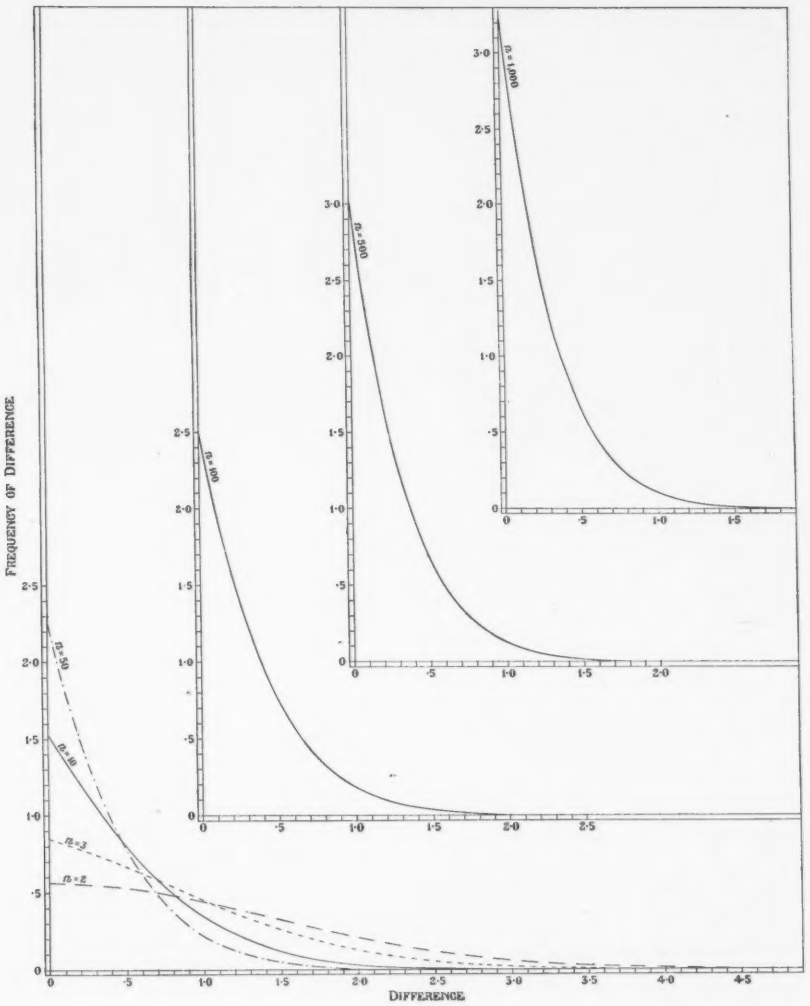


Diagram 1

APPROXIMATE FREQUENCY DISTRIBUTIONS
OF DIFFERENCES BETWEEN SECOND & THIRD INDIVIDUALS

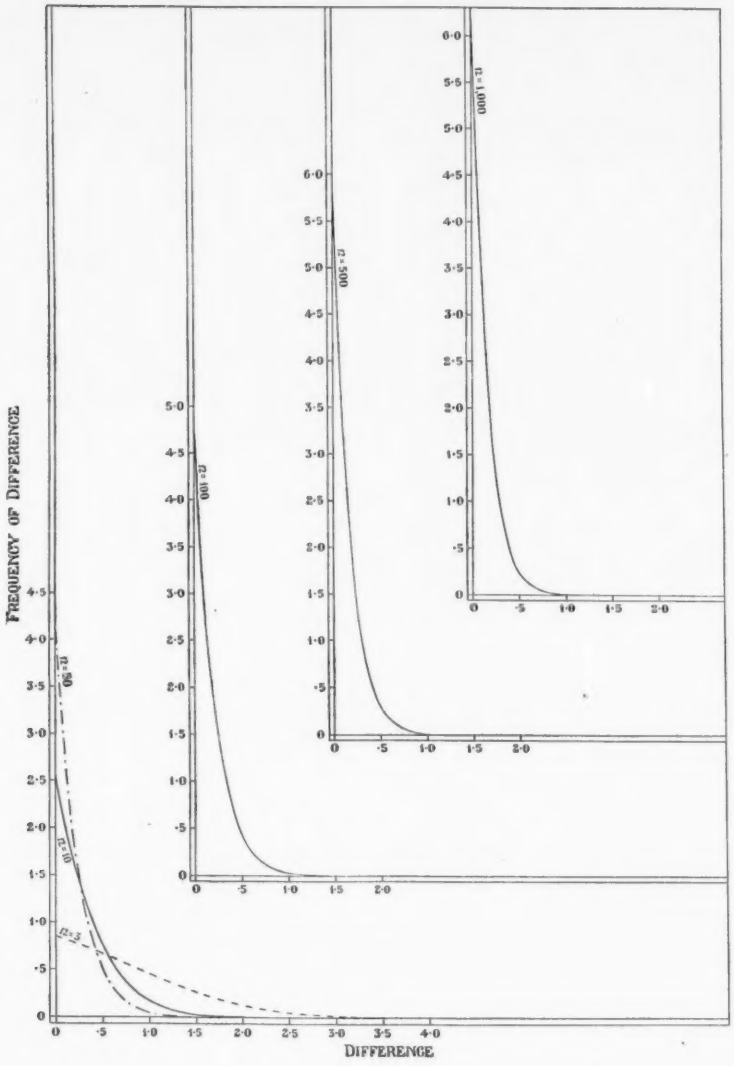


Diagram 2

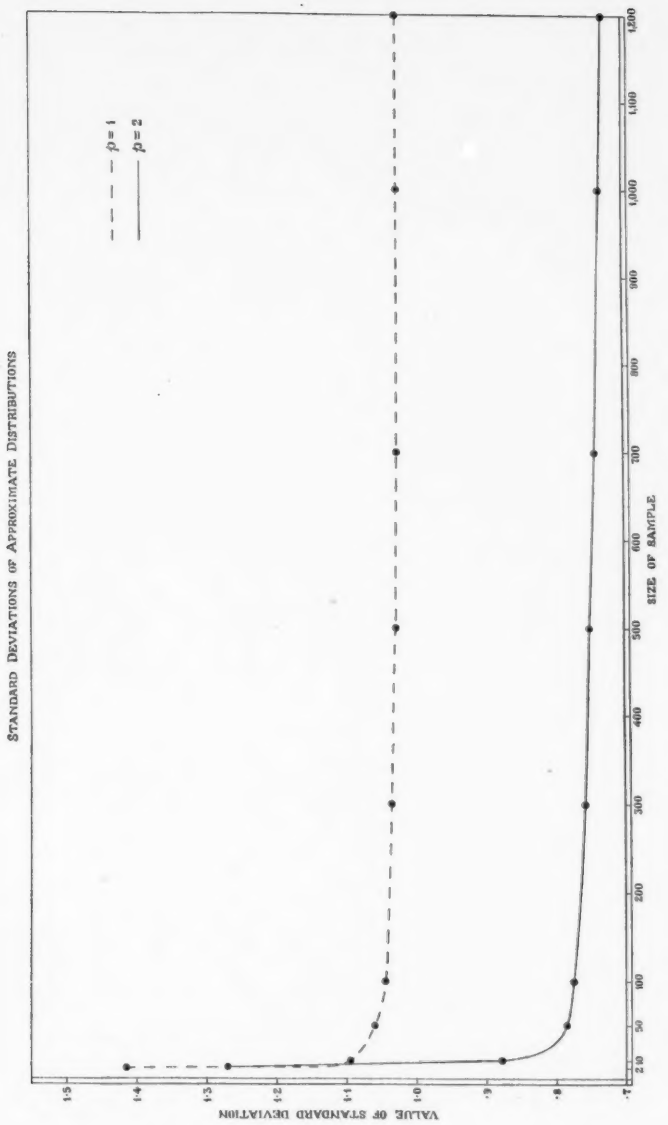

STANDARD DEVIATIONS OF APPROXIMATE DISTRIBUTIONS
p = 1
p = 2
VALUE OF STANDARD DEVIATION
1·5
1·4
1·3
1·2
1·1
1·0
·9
·8
·7
SIZE OF SAMPLE
20
50
100
200
300
400
500
600
800
900
1,000
1,100
1,200

Diagram 3

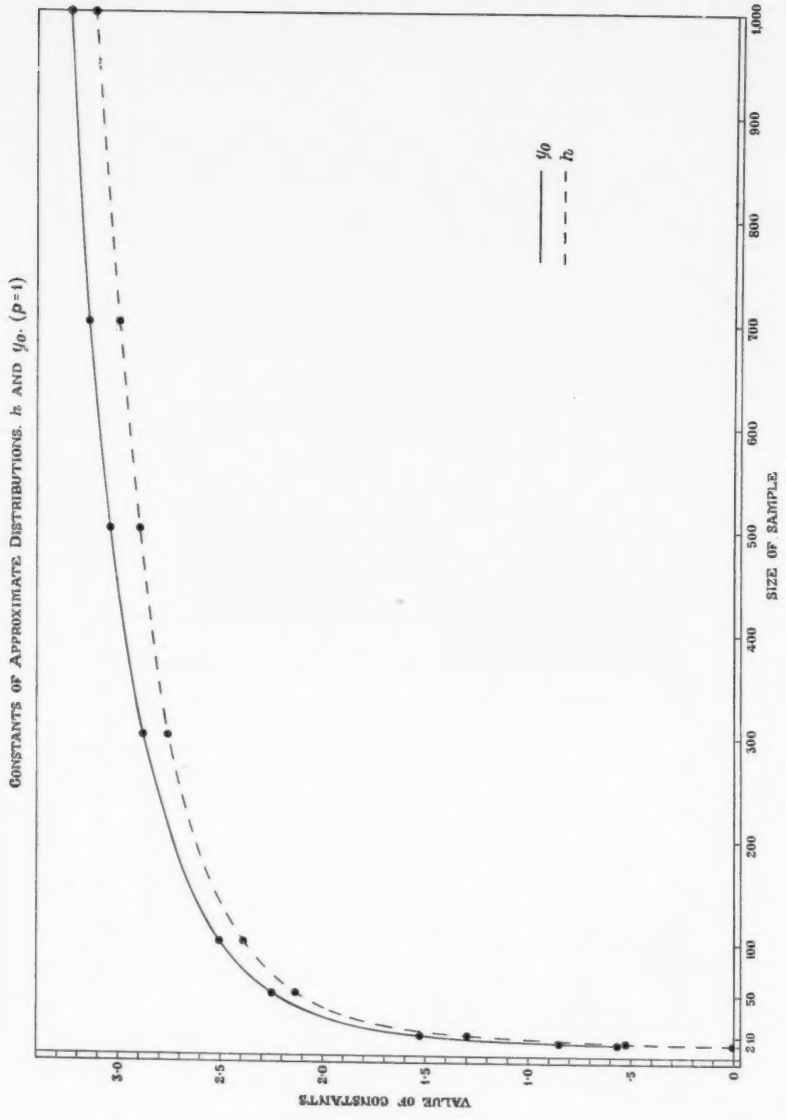

CONSTANTS OF APPROXIMATE DISTRIBUTIONS. h AND y0. (p=1)
y0
h
SIZE OF SAMPLE
VALUE OF CONSTANTS
1,000
900
800
700
600
500
400
300
200
100
50
10
3·0
2·5
2·0
1·5
1·0
·5
0

Diagram 4

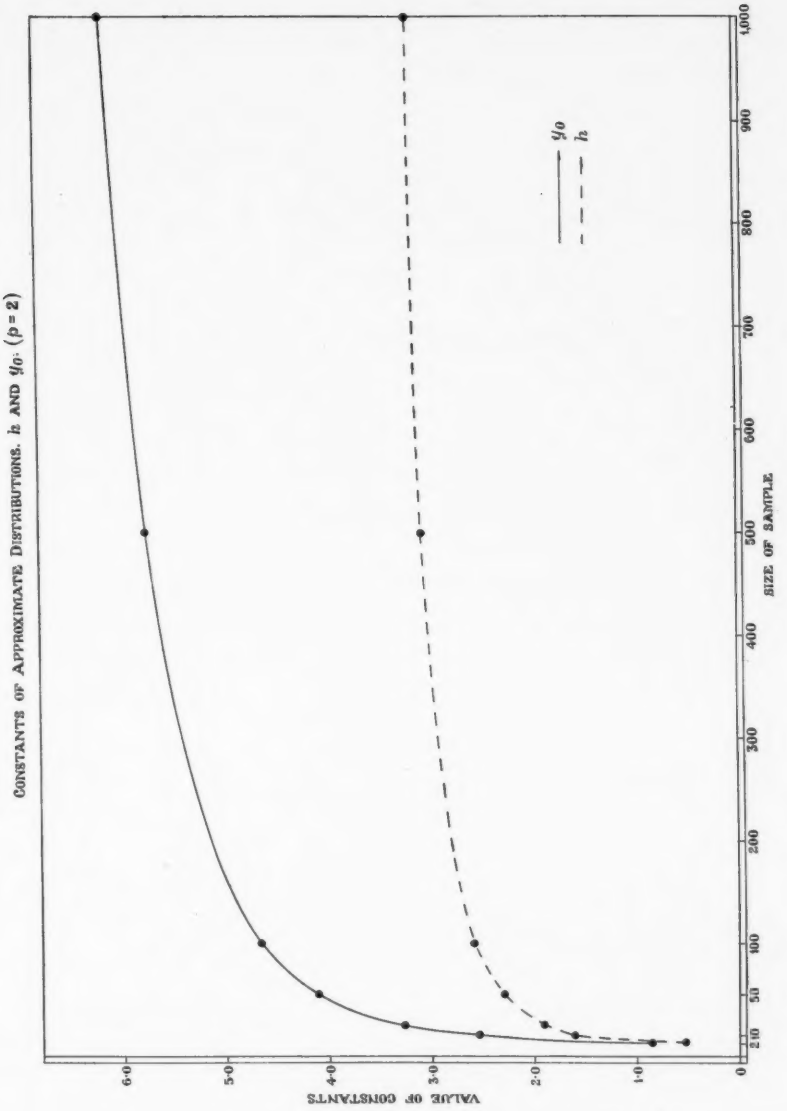

CONSTANTS OF APPROXIMATE DISTRIBUTIONS. h AND y_0. ($p = 2$)
y_0
h
VALUE OF CONSTANTS
SIZE OF SAMPLE
0
1·0
2·0
3·0
4·0
5·0
6·0
10
50
100
200
300
400
500
600
700
800
900
1,000

Diagram 5

Let $\chi_{2,1}$ be the mean of the differences, all taken positively, between pairs of individuals drawn at random from the original population, $\chi'_{2,1}$ the mean of the differences, all taken positively, between pairs of brothers drawn at random.

Then from the previous theory

$$\chi_{2,1} = \frac{2}{\sqrt{\pi}}\sigma \quad \text{.....(xxvii)*,}$$

σ being the standard deviation of the original population and we have to find the mean value of $\chi'_{2,1}$.

Let the population of pairs of brothers be represented by the normal frequency surface

$$z = \frac{1}{2\pi\sigma^2\sqrt{1-r^2}}\, e^{-\frac{1}{2}\frac{1}{1-r^2}\left(\frac{x^2-2rxy+y^2}{\sigma^2}\right)}$$

and let us determine the mean value of $|x-y|$.

Let $$x = \frac{X-Y}{\sqrt{2}}, \quad y = \frac{X+Y}{\sqrt{2}},$$

then $$X = \frac{x+y}{\sqrt{2}}, \quad Y = \frac{x-y}{\sqrt{2}}$$

and $$|x-y| = |\sqrt{2}Y| = |Y'| \text{ say.}$$

The surface may then be written

$$\frac{1}{2\pi\sigma^2\sqrt{1-r^2}}\, e^{-\frac{1}{2}\frac{1}{(1-r^2)\sigma^2}[X^2+Y^2-r(X^2-Y^2)]} = \frac{1}{2\pi\sigma^2\sqrt{1-r^2}}\, e^{-\frac{1}{2}\frac{X^2}{(1+r)\sigma^2}}\, e^{-\frac{1}{2}\frac{Y^2}{(1-r)\sigma^2}}.$$

The frequency distribution for Y may be obtained by integrating for X between $-\infty$ and ∞ namely:

$$z = \frac{1}{\sigma\sqrt{2\pi(1-r)}}\, e^{-\frac{1}{2}\frac{Y^2}{(\sigma\sqrt{1-r})^2}},$$

whence the frequency distribution for Y' is

$$z = \frac{1}{\sqrt{2\pi}\sigma\sqrt{2(1-r)}}\, e^{-\frac{1}{2}\frac{Y'^2}{[\sigma\sqrt{2(1-r)}]^2}} \quad \text{.....(xxviii).}$$

Therefore $$\chi'_{2,1} = \text{mean}\,|Y'| = \sqrt{\frac{2}{\pi}}\,\sigma\sqrt{2(1-r)} = \frac{2}{\sqrt{\pi}}\,\sigma\sqrt{1-r} \quad \text{.....(xxix)†,}$$

and $$\frac{\chi'_{2,1}}{\chi_{2,1}} = \sqrt{1-r} \quad \text{.....(xxx).}$$

* Cf. footnote, p. 101.

† This formula is a particular case for $\sigma_1 = \sigma_2$ of one first obtained by Pearson in the form

$$\frac{S_{x>y}\left\{\frac{x}{\sigma_1} - \frac{y}{\sigma_2}\right\}}{N} = \sqrt{\frac{1-r}{\pi}},$$

Drapers' Company Research Memoirs, No. IV. It can be used to determine r if σ_1, σ_2 are known, or are obtained from the observations. The method here proposed differs from Pearson's, by the fact that we eliminate the unknown standard deviation of the original population by taking the ratio of the mean of our two sets of differences.

It is worth while finding the probable error of r calculated from this formula. We shall obtain a first approximation only, assuming statistical differentials may be treated as mathematical differentials and that $\chi'_{2,1}, \chi_{2,1}$ are each obtained from averaging N pairs. Then

$$\left.\begin{aligned}\sigma^2\chi_{2,1} &= \frac{\sigma^2 d_{2,1}}{N}\\ \sigma^2\chi'_{2,1} &= \frac{\sigma^2 d'_{2,1}}{N}\end{aligned}\right\} \quad \text{(xxxi)},$$

where $d_{2,1}$ denotes the difference, taken positively, between a pair drawn at random; $d'_{2,1}$ the difference, taken positively, between a pair of brothers.

Taking logarithmic differentials of (xxx)

$$-\frac{1}{2}\frac{\delta r}{(1-r)} = \frac{\delta\chi'_{2,1}}{\chi'_{2,1}} - \frac{\delta\chi_{2,1}}{\chi_{2,1}},$$

$$\frac{1}{4}\frac{(\delta r)^2}{(1-r)^2} = \frac{(\delta\chi'_{2,1})^2}{\chi'^2_{2,1}} + \frac{(\delta\chi_{2,1})^2}{\chi^2_{2,1}} - \frac{2\delta\chi_{2,1}\delta\chi'_{2,1}}{\chi_{2,1}\chi'_{2,1}}.$$

Summing for all possible samples and dividing by their number

$$\frac{1}{4}\frac{\sigma_r^2}{(1-r)^2} = \frac{\sigma^2\chi'_{2,1}}{\chi'^2_{2,1}} + \frac{\sigma^2\chi_{2,1}}{\chi^2_{2,1}} - \frac{2[\delta\chi_{2,1}\delta\chi'_{2,1}]}{\chi_{2,1}\chi'_{2,1}},$$

where the symbol [ ] denotes mean value.

But $$\sigma^2\chi_{2,1} = \frac{\sigma^2 d_{2,1}}{N} = \frac{1}{N}\left\{2\left(1-\frac{2}{\pi}\right)\right\}\sigma^2 \quad \text{.....(xxxii)*}$$

and $$[\delta\chi_{2,1}\delta\chi'_{2,1}] = 0,$$

because we are dealing with two independent processes of sampling.

It remains to find $\sigma^2\chi'_{2,1}$.

$$\sigma^2\chi'_{2,1} = \frac{\sigma^2 d'_{2,1}}{N} = \frac{1}{N}\sigma^2|Y'|$$

$$= \frac{1}{N}\left\{2\int_0^\infty \frac{Y'^2 e^{-\frac{1}{2}\frac{Y'^2}{\{\sigma\sqrt{2(1-r)}\}^2}} dY'}{\sqrt{2\pi}\,\sigma\sqrt{2(1-r)}} - \chi'^2_{2,1}\right\}$$

$$= \frac{2}{N}\sigma^2(1-r) - \frac{1}{N}\frac{4}{\pi}\sigma^2(1-r)$$

$$= \frac{2\sigma^2(1-r)}{N}\left\{1-\frac{2}{\pi}\right\} \quad \text{.....(xxxiii).}$$

Therefore $$\frac{\sigma^2\chi'_{2,1}}{\chi'^2_{2,1}} = \frac{\frac{1}{2}(\pi-2)}{N} = \frac{\sigma^2\chi_{2,1}}{\chi^2_{2,1}}$$

and $$\frac{\sigma_r^2}{(1-r)^2} = 4\left\{\frac{\sigma^2\chi'_{2,1}}{\chi'^2_{2,1}} + \frac{\sigma^2\chi_{2,1}}{\chi^2_{2,1}}\right\} = \frac{4(\pi-2)}{N},$$

or $$\sigma_r = 2(1-r)\sqrt{\frac{\pi-2}{N}} = 2{\cdot}137\,\frac{(1-r)}{\sqrt{N}} \quad \text{.....(xxxiv),}$$

* Cf. footnote, p. 111.

which gives the following ratios to the probable errors as they would be obtained by the product-moment method with the same value of N.

r	Ratio	r	Ratio
·1	1·943	·6	1·336
·2	1·781	·7	1·257
·3	1·644	·8	1·187
·4	1·526	·9	1·125
·5	1·425	1·0	1·068

Thus the probable error is always larger than that given by the product-moment method, but if a large enough number of pairs be taken there seems no reason why the method should not be practicable.

(3) In order that we might be able to obtain from the calculated r the probability of a deviation in excess or defect, we should really require to know not only the standard error but also the frequency distribution of r calculated in this way. This seems a very difficult problem. For large values of N it would presumably be approximately normal, but for smaller values it would seem to depend on the frequency distribution of differences. A knowledge of the frequency distribution of differences, therefore, may well be a first step in its solution.

At present our knowledge of the frequency distribution of differences is complete only for samples of two or three individuals, though good approximations have been obtained for larger samples. This together with consideration of probable errors suggests that it is for large numbers of small samples that Galton's Difference Problem is most likely to lead to practical results.

But it must be acknowledged, in conclusion, that the complicated theory of Galton's Difference Method makes it evident that very considerable difficulties will attend its application in practice, if an intelligible and well defined meaning is to be attached to the results obtained.

I am indebted to Professor Pearson for suggesting the subject of this investigation and for some valuable suggestions and criticism, and to Miss Ida M^{c}Learn for the preparation of the diagrams.

ON THE MULTIPLE CORRELATION OF BROTHERS, BEING A NOTE ON MR J. O. IRWIN'S MEMOIR, AND ON MY STATEMENT OF THE APPLICATION OF GALTON'S DIFFERENCE PROBLEM TO THE DETERMINATION OF THE DEGREE OF RELATIONSHIP OF BROTHERS, MADE IN AUGUST 1902. (*Biometrika*, Vol. I. p. 399.)

By KARL PEARSON, F.R.S.

Mr Irwin's paper does not seem to me to put the case for the application of Galton's Difference Problem to the subject of brothers adequately. I had not at the date his paper went to press examined my statement of 1902, and took on faith that what he wrote in his paper (ftn. p. 101) was in accordance with my remark of more than twenty years ago. Returning to that paper I find no ground for his interpretation of my statement. My words are the following:

> Anyhow I think we may say that if the individuals [i.e. a *pair*] be taken at random from a population, then the probable value* of the standard deviation of that population is nearly $\frac{33}{46}$ of the difference between the two individuals. Thus by averaging the differences between pairs of individuals taken at random we can obtain fairly readily an appreciation of the standard deviation, i.e. the variability of the general population. Further, if we take individuals not quite at random, but from correlated groups, e.g. pairs of brothers selected at random, the $\frac{33}{24}$th of the average difference of the pairs will be the standard deviation of the correlated groups, e.g. a group of brethren; hence the degree of relationship between such correlated individuals may be determined. This is only a suggestion of one of the many possible uses of Galton's difference problem. (*Biometrika*, Vol. I. p. 399.)

The paragraph may be somewhat obscurely worded, but I think the suggestion is definite in the fact that the pair has to be taken from a *group of brethren*, i.e. a group all of whom are offspring of the same parental pair. Mr Irwin supposes me to say that the two individuals are to be taken from a group all of whom have a brother of a given intensity of character, i.e. from the array of a fraternal correlation table. My statement says a "group of brethren" and thus interpreted there appears to me to be no error in it. A group of brethren differs from the general population in having all its ancestry right back the same. If σ be the standard deviation of the general population, then the standard deviation of such a group of brethren is $\sigma\sqrt{1-R^2}$, where R is the multiple correlation coefficient of the son on all his ancestors. If we pick out pairs of individuals at random from a group of brothers, and from the general population, then the ratio of the mean differences of such respective pairs is such that:

$$\frac{\text{Mean difference of fraternal pairs}}{\text{Mean difference of general population pairs}} = 1 - R^2.$$

[* The use of the term "probable value" in the light of our present knowledge is incorrect. The statement as to averaging differences is correct. But the mean or average value is not the probable value, for the distribution is known to be skew, as is well illustrated in Mr Irwin's paper.]

It is thus possible to find R by this process, and R involves at once a determination of r the so-called "correlation of brothers" as found by a fraternal correlation table.

The difficulty of the process in the case of *human* families is their small size; this renders it impossible to draw pairs from an adequately large group of brothers, who would have $\sigma\sqrt{1-R^2}$ for their standard deviation. But this difficulty (which does not exist in the case of some other species) can be got over as follows:

Suppose we take samples of size n from a population of standard deviation Σ following the normal law of distribution. Then the distribution curve of the standard deviation s in the samples is known to be

$$z = \text{const.} \times \frac{s^{n-2}}{\Sigma^{n-1}} e^{-\frac{ns^2}{2\Sigma^2}},$$

and the chance of a standard deviation falling between s and $s + ds$ is proportional to

$$\frac{s^{n-2}}{\Sigma^{n-1}} e^{-\frac{ns^2}{2\Sigma^2}} ds.$$

But this must also be the chance of the square of the standard deviation s^2 or μ_2 having a value between s^2 and $(s + \delta s)^2$, or between μ_2 and $\mu_2 + \delta\mu_2$. Expressing the result in terms of μ_2, we have the chance proportional to

$$\left(\frac{\mu_2}{\Sigma^2}\right)^{\frac{n-3}{2}} e^{-\frac{n\mu_2}{2\Sigma^2}} \frac{d\mu_2}{\Sigma^2},$$

or the frequency curve for the distribution of μ_2 is

$$z_y = \frac{z_0}{\Sigma^2}\left(\frac{\mu_2}{\Sigma^2}\right)^{\frac{n-3}{2}} e^{-\frac{n\mu_2}{2\Sigma^2}}.$$

The mean value of μ_2 in samples, or $\bar{\mu}_2$, is accordingly

$$\bar{\mu}_2 = \Sigma^2 \frac{\int_0^\infty \left(\frac{\mu_2}{\Sigma^2}\right)^{\frac{n-3}{2}} \frac{\mu_2}{\Sigma^2} e^{-\frac{n\mu_2}{2\Sigma^2}} \frac{d\mu_2}{\Sigma^2}}{\int_0^\infty \left(\frac{\mu_2}{\Sigma^2}\right)^{\frac{n-3}{2}} e^{-\frac{n\mu_2}{2\Sigma^2}} \frac{d\mu_2}{\Sigma^2}}$$

$$= \frac{2\Sigma^2}{n} \frac{\int_0^\infty z^{\frac{1}{2}(n-1)} e^{-z} dz}{\int_0^\infty z^{\frac{1}{2}(n-3)} e^{-z} dz} = \frac{2\Sigma^2}{n} \frac{\Gamma\left(\frac{n+1}{2}\right)}{\Gamma\left(\frac{n-1}{2}\right)}$$

$$= \frac{n-1}{n} \Sigma^2.$$

Hence the mean standard deviation squared of *pairs* of brothers is obtained by taking $n = 2$, or is equal to $\frac{1}{2}\Sigma^2 = \frac{1}{2}\sigma^2(1 - R^2)$.

Now let x_1 and x_2 be two individuals, then their mean $= \frac{1}{2}(x_1 + x_2)$ and their standard deviation squared equals

$$\frac{1}{2}\{(x_1 - \frac{1}{2}(x_1 + x_2)) + (x_2 - \frac{1}{2}(x_1 + x_2))^2\}$$
$$= \frac{1}{4}(x_1 - x_2)^2.$$

Accordingly we want the mean value of $\frac{1}{4}(x_1 - x_2)^2$ for all pairs, say N, and this equals

$$\frac{\frac{1}{4} S\{x_1 - \bar{x} - (x_2 - \bar{x})\}^2}{N}$$
$$= \frac{1}{4}(\sigma_1^2 + \sigma_2^2 - 2r\sigma_1\sigma_2),$$

where r is the so-called correlation of brothers and σ_1 is the standard deviation of the first, σ_2 of the second brother. Taking these equal, we have both equal σ; and the mean standard deviation squared of pairs of brothers $= \frac{1}{2}\sigma^2(1-r) = \frac{1}{2}\sigma^2(1 - R^2)$. Hence $r = R^2$, or the correlation of brothers is equal to the squared multiple correlation coefficient of the offspring on all ancestry. A knowledge of R gives r; and Mr Irwin translates my $1 - R^2$ into $1 - r^2$ and supposes I meant pairs from an array and not pairs from a group of brothers as I wrote.

We may now examine the matter from another aspect. Consider the variate x_0 in correlation with the variates $x_1, x_2, \ldots x_n$, the correlation coefficients being given by r_{uv} between the uth and vth variates. Then it is well known that the multiple correlation coefficient of x_0 on $x_1, x_2, \ldots x_n$ is given by R_n, where

$$R_n^2 = 1 - \frac{\Delta}{\Delta_{00}},$$

Δ being the determinant

$$\Delta = \begin{vmatrix} 1 & r_{01} & r_{02} & \ldots r_{0n} \\ r_{10} & 1 & r_{12} & \ldots r_{1n} \\ r_{20} & r_{21} & 1 & \ldots r_{2n} \\ \ldots\ldots & \ldots\ldots & \ldots\ldots & \ldots\ldots \\ r_{n0} & r_{n1} & r_{n2} & \ldots 1 \end{vmatrix},$$

and Δ_{00} the minor of the leading constituent. If Δ takes the form

$$\begin{vmatrix} 1, & r, & r, & \ldots r \\ r, & 1, & \rho, & \ldots \rho \\ r, & \rho, & 1, & \ldots \rho \\ \ldots\ldots & \ldots\ldots & \ldots\ldots & \ldots\ldots \\ r, & \rho, & \rho, & \ldots 1 \end{vmatrix}$$

it is easy to show that

$$R_n = r\sqrt{\frac{n}{1 + (n-1)\rho}}.$$

Accordingly if $\rho = r$ and n be made indefinitely great, we have

$$R_\infty = r\sqrt{\frac{1}{\rho}} = \sqrt{r}.$$

The form thus taken by R is, however, the form suitable to an indefinitely great group of brothers. We learn accordingly that:

(*a*) The multiple correlation coefficient of the offspring on an indefinitely extended ancestry is equal to the multiple correlation coefficient of one brother on an indefinitely great number of brothers.

(*b*) Either of these multiple correlation coefficients is the square root of the ordinary correlation coefficient for two brothers.

(*c*) The simple correlation coefficient for two brothers is numerically of about the order 0·5; it follows that the multiple fraternal correlation is of an order somewhat over 0·7. It is this multiple fraternal correlation coefficient which is directly determined, when we take the ratio of the mean difference between pairs of brothers to the mean difference between pairs of individuals from the general population.

Further, the importance of the multiple fraternal correlation (R), if determined, is very great, for on the assumption of the Law of Ancestral Heredity—i.e. the decadence of individual ancestral correlations in geometric progression—it enables us to write down with comparative ease the expression for the mean value of the offspring with a given ancestry. The fundamental formulae for the Law of Ancestral Heredity are given in *Biometrika*, Vol. VIII. pp. 239—240, and may be briefly recapitulated here. x_0 is the character value in the individual offspring in the last generation. X_p is the character of the "mid-parent" in the pth ancestral generation, $\bar{x}_0$ and $\bar{X}_p$ are the corresponding means and σ_0, Σ_p the corresponding standard deviations; σ_p the standard deviation of individuals in the pth ancestral generation; ρ_{p-q} the *average* correlation of individual ancestors in the pth generation with their offspring in the qth generation ($p > q$); ϵ_p the coefficient of assortative mating in the parents of the pth ancestral generation, ρ_{p-q} the *average* coefficient of correlation between individual ancestors of the pth and individual offspring of the qth ancestral generation, and r_{p-q} the correlation between the mid-parent of the pth and the mid-parent of the qth generation. Then the multiple regression formula of the offspring on the series of ancestral mid-parents being

$$x_0 - \bar{x}_0 = S\left\{c_p \frac{\sigma_0}{\Sigma_p}(X_p - \bar{X}_p)\right\},$$

with a variability

$$\Sigma_0 = \sigma_0 \sqrt{1 - S(c_p r_{p0})},$$

and the geometrical decrease of ρ_{p-q} being assumed, it is shown that

$$\Sigma_p = \frac{\sigma_p}{(\sqrt{2})^p}\sqrt{1+\epsilon_p},$$

$$r_{p-q} = (\sqrt{2})^{p-q}\,\rho_{p-q}/\sqrt{(1+\epsilon_p)(1+\epsilon_q)}.$$

Accordingly if assortative mating be constant throughout the generations, r_{p-q} follows the geometric law if ρ_{p-q} does. Then taking

$$r_{p-q} = \alpha\beta^{p-q}, \quad c_p = \gamma\eta^p,$$

we find
$$\beta - \eta = \gamma\eta,$$
$$\eta^2 - \eta\frac{(1-2\alpha\beta^2+\beta^2)}{\beta(1-\alpha)}+1=0,$$
which give γ and η in terms of α and β. Further,
$$\Sigma_0 = \sigma_0\sqrt{1-\frac{\gamma\eta\alpha\beta}{1-\eta\beta}} = \sigma_0\sqrt{1-R^2},$$
or
$$R^2 = \frac{\gamma\eta\alpha\beta}{1-\eta\beta} = r,$$
the fraternal correlation as we have shown on p. 131 above.

Accordingly a knowledge of R^2 and the equations for η and γ above enable us to find β, γ and η in terms of α or $\alpha\beta$. But for parental correlation we have
$$\alpha\beta = r_1 = \sqrt{2}\,\rho_1/(1+\epsilon),$$
or $\alpha\beta$ is known in terms of the mean (ρ_1) of the two parental correlations and the coefficient ϵ of assortative mating.

We will write
$$\rho_1/(1+\epsilon) = \rho_1'.$$

Substituting for $\gamma\eta$ in the value for R^2, we have
$$R^2 = \frac{(\beta-\eta)\,\alpha\beta}{1-\eta\beta},$$
whence we find
$$\eta = \frac{R^2-\alpha\beta^2}{\beta(R^2-\alpha)} = \frac{R^2-\beta\sqrt{2}\,\rho_1'}{\beta R^2-\sqrt{2}\,\rho_1'}.$$

Substituting this value of η in the quadratic for η above, we find
$$\beta^2 - 2\beta\sqrt{2}\,\rho_1' + \frac{2\rho_1'^2 - R^2(1-R^2)}{R^2} = 0,$$
or
$$\beta = \sqrt{2}\,\rho_1' \pm \sqrt{\frac{(1-R^2)(R^2-2\rho_1'^2)}{R^2}}.$$

We shall consider the alternate roots later.

Then
$$\alpha = \sqrt{2}\,\rho_1'/\beta,$$
and
$$\rho_{p-q} = \rho_1\left(\frac{\beta}{\sqrt{2}}\right)^{p-q-1}$$
gives the geometric law of degradation of the average individual ancestral correlations. Finally
$$\gamma\eta = \frac{(\beta^2-1)\,R^2}{\beta R^2-\sqrt{2}\,\rho_1'} \text{ and } \eta = \frac{R^2-\beta\sqrt{2}\,\rho_1'}{\beta R^2-\sqrt{2}\,\rho_1'}$$
enable us to determine
$$c_p = \gamma\eta\,.\,\eta^{p-1}$$
and complete the solution for the multiple regression equation of offspring on ancestry.

Thus, a knowledge of:

(*a*) Parental correlat.ons;

(*b*) Fraternal correlations;

(*c*) Coefficient of assortative mating, and the assumptions

(i) that the population is stable

$$\sigma_p = \sigma_0, \text{ for all values of } p,$$
$$\epsilon_p = \epsilon \quad ,, \quad ,,$$

(ii) that the individual ancestral correlations follow a geometric law of degradation

enable us to determine the multiple regression equation, without determining the individual ancestral correlations. There is, however, an exception which is so important that it needs special consideration. Our equations become indeterminate when $\beta = 1$. We will deal with this case.

Case of equality of all coefficients of correlation. When $\beta = 1$, $r_{p-q} = \alpha$, whatever the values of p and q. All the mid-parents have the same correlations with each other and with any generation of offspring. We therefore need the multiple regression equation, when the fundamental determinant is

$$\Delta = \begin{vmatrix} 1, & r, & r, \ldots & r \\ r, & 1, & r, \ldots & r \\ r, & r, & 1, \ldots & r \\ \ldots\ldots & \ldots\ldots & \ldots\ldots & \ldots\ldots \\ r, & r, & r, \ldots & 1 \end{vmatrix} \text{ to } n+1 \text{ rows and columns.}$$

Add up the last n rows, divide their sums by $\dfrac{1+\overline{n-1}\,r}{r}$ and subtract from the first row. We have

$$\Delta = \begin{vmatrix} 1 - \dfrac{nr^2}{1+n-1\,r}, & 0, & 0, \ldots & 0 \\ -r & , 1, & r, \ldots & r \\ r & , r, & 1, \ldots & r \\ \ldots\ldots\ldots\ldots & \ldots\ldots\ldots\ldots & \ldots\ldots & \ldots\ldots \\ r & , r, & r, \ldots & 1 \end{vmatrix},$$

or
$$\Delta = \left(1 - \frac{nr^2}{1+(n-1)\,r}\right), \quad \Delta_{00} = \frac{(1-r)(1+nr)}{1+(n-1)\,r}\,\Delta_{00}.$$

Similarly
$$\Delta_{00} = \frac{(1-r)\,(1+(n-1)\,r)}{1+(n-2)\,r}\,\Delta_{0011},$$

where Δ_{0011} is the minor obtained by cancelling the first two rows and first two columns.

Now let us endeavour to find

$$\Delta_{01} = -\begin{vmatrix} r, & r, & r, \ldots & r \\ r, & 1, & r, \ldots & r \\ r, & r, & 1, \ldots & r \\ \ldots\ldots & \ldots\ldots & \ldots\ldots & \ldots\ldots \\ r, & r, & r, \ldots & 1 \end{vmatrix}.$$

Add up the last $n-1$ rows, divide their sums by $\dfrac{1+(n-2)r}{r}$ and subtract from the first row. We have

$$\Delta_{01} = -\begin{vmatrix} r - \dfrac{(n-1)r^2}{1+(n-2)r}, & 0, & 0, \ldots 0 \\ r, & 1, & r, \ldots r \\ r, & r, & 1, \ldots r \\ \cdots\cdots & \cdots & \cdots\cdots \\ r, & r, & r, \ldots 1 \end{vmatrix}$$

$$= -\left(r - \frac{(n-1)r^2}{1+(n-2)r}\right)\Delta_{0011}$$

$$= -\frac{r(1-r)}{1+(n-2)r}\,\Delta_{0011}.$$

Substituting the value already found for Δ_{0011}, we have

$$\frac{\Delta_{01}}{\Delta_{00}} = -\frac{r}{1+(n-1)r} = \frac{\Delta_{0p}}{\Delta_{00}}$$

by the symmetry of the determinant.

Accordingly the multiple regression equation for the mean value of an array of x_0's, when determined from a series of variates $x_1, x_2, \ldots x_n$ of standard deviations $\sigma_1, \sigma_2, \ldots \sigma_n$, when all the correlations are the same, is

$$\tilde{x}_0 - \bar{x}_0 = \sigma_0 \sum_1^n \frac{r}{1+(n-1)r}\,\frac{x_p - \bar{x}_p}{\sigma_p},$$

where $\Sigma_0 = \sigma_0\sqrt{1-R^2}$ gives the variability, and $R^2 = \dfrac{nr^2}{1+(n-1)r}$ is the multiple correlation coefficient *.

When n becomes indefinitely large $R^2 = r$ and the regression coefficient becomes the sum of an infinitely large series of infinitely small terms. That this series will take a finite limit is clear; for example, if we take $(x_p - \bar{x}_p)/\sigma_p = \text{constant} = q$, then $\tilde{x}_0 - \bar{x}_0 = \sigma_0 \dfrac{rnq}{1+(n-1)r} = q\sigma_0$ in the limit.

Returning to the case of predicting the probable value of a character from a knowledge of the ancestral values, we have $r = r_{p-q} = \alpha\beta^{p-q} = \alpha = \sqrt{2}\,\rho_1'$. Thus

$$\tilde{x} - \bar{x}_0 = \sigma_0 \frac{\sqrt{2}\,\rho_1'}{1+(n-1)\sqrt{2}\,\rho_1'} \sum_1^n \frac{X_p - \bar{X}_p}{\Sigma_p},$$

$$\Sigma_0 = \sigma_0\sqrt{1-R^2}, \text{ where } R^2 = \frac{2n\rho_1'^2}{1+\sqrt{2}\,(n-1)\,\rho_1'}.$$

The previous remark holds good here, that as n becomes very large we have an indefinitely large number of indefinitely small terms. But there is no advantage

* I do not remember publishing this result before, but have given it occasionally in lecture to illustrate multiple correlation.

in taking n indefinitely large. When $n = 10$, excluding cousin marriages, there will be 2048 ancestors in the 10th ancestral generation and the "mid-parent" provided by them will be essentially equal to the mean of that generation or $X_p = \bar{X}_p$.

We may note that in this case

$$\rho_{p-q} = \rho_1 \left(\frac{\beta}{\sqrt{2}}\right)^{p-q-1} = \rho_1 \left(\frac{1}{\sqrt{2}}\right)^{p-q-1},$$

or the radix of the geometrical progression is ·707107 *.

If we take our system of equal correlations, the deduction of the probable character of one brother from n others, then $\sigma_0 = \sigma_p$ for all values of p, and if r be the simple fraternal correlation

$$\tilde{x}_0 - \bar{x}_0 = \frac{nr}{1 + (n-1)\,r}\left(\frac{S(x_p)}{n} - \bar{x}_0\right),$$

$$\Sigma_0 = \sigma_0 \sqrt{1 - R'^2},$$

where

$$R'^2 = \frac{nr^2}{1 + (n-1)\,r}.$$

When n becomes infinite we have the array of brothers or sisters, which *could* form a family. But this must be the array of offspring from the same total ancestry, i.e. for $n = \infty$,

$$R'^2 = R^2,$$

or

$$r = \sqrt{2}\,\rho_1'.$$

Here $\rho_1' = \rho_1/(1 + \epsilon)$.

Taking stature, forearm and span, we find the mean value of ϵ is ·235667†. The mean parental inheritance for the same data was ·45925‡. Hence $\rho_1' = ·371{,}6616$ and $\sqrt{2}\,\rho_1' = ·5256$. The mean value of r the correlation coefficient in the corresponding tables for stature, forearm and span in pairs of brothers and in pairs of sisters was ·5250§, which is sensibly in accord with the above result. The advantage of such a formula as that just given is that it enables us to deduce fraternal from parental and assortative mating correlations, or *vice versa*. The formula, however, essentially depends on β being nearly equal to unity, a result which would flow from the hypothesis that the associations of the mid-parents of each generation of ancestry to each other and to the offspring are of the same intensity. I now propose to develop the formulae provided in this note numerically from the data given by my family measurements of 1893. The cards then issued were limited to four adult siblings in each family, and so in many cases it was only possible to measure one brother and one sister. Thus the pairs of brothers and the pairs of sisters were not as numerous as we could have wished, while we had ample material for parental inheritance and for brother-sister pairs. To illustrate points in this

* For example the mean parental correlation coefficient being taken at ·460 (*Biometrika*, Vol. II. p. 378) we have ·3253, ·2300, ·1626, ·1150, ·0813, etc. for the successive grandparental average correlations. These are by no means out of accord with our somewhat limited experience.

† *Biometrika*, Vol. II. p. 373.

‡ *Ibid.* p. 378.

§ *Ibid.* p. 387.

present note and in Mr Irwin's preceding paper, I have confined myself to the single character of stature in pairs of brothers and pairs of sisters. Table I gives the differences in stature of 341 pairs of brothers and 482 pairs of sisters arranged in $\frac{1}{4}''$ units*. I first took the differences given in this table, squared them and took their weighted mean. This had to be divided by 64, because the units are in quarter inches and because what we need (see p. 131) is the mean value of $\frac{1}{4}(x_1-x_2)^2$. In this way I found

$$\tfrac{1}{2}\sigma^2(1-R^2) = 1{\cdot}557{,}345, \text{ for brothers,}$$

$$= 1{\cdot}597{,}267, \text{ for sisters.}$$

But σ^2 in inches was 6·85866 for 682 brothers and 6·94701 for 964 sisters.

Calculating R^2, from the above values, I found

For Brothers: $R^2 = {\cdot}545{,}875$; for Sisters: $R^2 = 540{,}157$.

Accordingly, $R = {\cdot}7388$ for brothers and $= {\cdot}7350$ for sisters; these are the values of the multiple correlation coefficients for an indefinitely large number of brothers and sisters respectively.

We can approach the subject from another side. The correlation coefficient for the 682 pairs of brothers was ·5196 and for the 964 pairs of sisters† was ·5370. Hence $R = \sqrt{r}$ gives us $R = {\cdot}7208$ and ·7328 for brothers and sisters respectively. These results are not bad, better in the case of sisters than of brothers; we must bear in mind, however, that our value for the mean squared standard deviation of pairs depends on selecting at random from a *normal* distribution.

TABLE I.

Distribution of Differences of Stature between Pairs of Brothers and Pairs of Sisters.

Frequency of Pairs.

Differences in quarter inches	*0*	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*	*11*	*12*	*13*
Brother-Pairs	11·5	24	29·5	17	35·5	18·5	24·5	22·5	31	15	18	10·5	18·5	5
Sister-Pairs	28	35	37	24	45	26	32	25	35·5	16	39	14·5	25·5	9
Differences in quarter inches	*14*	*15*	*16*	*17*	*18*	*19*	*20*	*21*	*22*	*23*	*24*	*25*	*26*	*27*
Brother-Pairs	10	5	11	6·5	5·5	5	3·5	2·5	3	—	3	2	1	—
Sister-Pairs	18·5	15	17	8	6	4	2	2	5	2	3	4	—	—
Differences in quarter inches	*28*	*29*	*30*	*31*	*32*	*33*	*34*	*35*	*36*	*37*	*38*	*39*	*40*	Totals
Brother-Pairs	1	—	—	—	—	—	1	—	—	—	—	—	1	341
Sister-Pairs	—	1	—	1	1	—	—	—	—	—	1	—	—	482

* I have to thank my colleague, Miss E. M. Elderton, for taking out this table from the original data cards and for finding several constants for me.

† The tables were rendered symmetrical, as there was no reason to distinguish for our present purpose between younger and elder siblings, so that they contained 682 and 964 pairs respectively.

We now proceeded to take the mean difference of pairs. This was found to be 2″·03391 for pairs of brothers and 2″·01686 for pairs of sisters, the two standard deviations being 2″·61890 and 2″·63572 respectively.

Calculating $r\,(=R^2)$ from the formula*

$$\left(\frac{\text{Mean difference of pairs}}{\text{S.D. of population}}\right)^2=\frac{4}{\pi}(1-r),$$

we have

$$r=\cdot526{,}289 \text{ for brothers, } =\cdot540{,}119 \text{ for sisters,}$$

the correlation table values being

$$r=\cdot5196 \text{ for brothers, } =\cdot5370 \text{ for sisters.}$$

The former values give

$$R=\cdot7255 \text{ for brothers and } \cdot7349 \text{ for sisters.}$$

Thus we have the following values of R the multiple fraternal correlation coefficient:

	Pairs of Brothers	Pairs of Sisters
From mean square S.D. of pairs ...	·7388	·7380
From correlation table value of r ...	·7208	·7328
From mean difference of brothers ...	·7255	·7349
Mean of three methods	·7284	·7342
Mean of results for brothers and sisters	·7311	
Corresponding to an average individual fraternal correlation	·5345	

It will be clear from this numerical investigation that the value of the multiple correlation coefficient can be found with reasonable accuracy by taking the squares of the differences of brothers and of non-brothers, and that

$$\frac{\text{Mean square difference of character in brothers}}{\text{Mean square difference of character in non-brothers}}=1-R^2.$$

Or, again we may use:

$$\frac{\text{Mean difference of character in brothers}}{\text{Mean difference of character in non-brothers}}=\sqrt{1-R^2}.$$

The simple correlation coefficient may be found from the multiple correlation coefficient by the relation $r=R^2$.

We may now proceed to determine the constants of the multiple regression formulae on the basis of the values of the multiple correlation coefficients determined from the mean square standard deviations of pairs, i.e.

$$R^2=\cdot545{,}875 \text{ for brothers, } =\cdot540{,}157 \text{ for sisters.}$$

We have first, however, to determine ρ_1 and ρ_1' for stature. Taking the mean of the two parental correlations for stature, these give

$$\rho_1=\cdot5040 \text{ for sons, and } =\cdot5085 \text{ for daughters.}$$

* "On Further Methods of measuring Correlation." *Biometric Laboratory Publications*, IV. p. 5.

Now the coefficient of assortative mating in stature for this same material is $\epsilon = \cdot 2804$. Accordingly

$$\rho_1' = \rho_1/(1+\epsilon) = \cdot 393{,}627 \text{ for sons, and } = \cdot 397{,}142 \text{ for daughters,}$$
$$\sqrt{2}\,\rho_1' = \cdot 556{,}673 \text{ for sons, and } \cdot 561{,}644 \text{ for daughters,}$$
$$2\rho_1'^2 = \cdot 309{,}885 \quad \text{,,} \quad \text{,,} \quad \cdot 315{,}444 \quad \text{,,} \quad \text{,,}$$

From these values we find from the quadratic (p. 133):

For Sons	For Daughters
$\beta = \cdot 999{,}759,$	$\beta = \cdot 999{,}024,$
or, $= \cdot 113{,}587,$	or, $= \cdot 124{,}264.$

Now the individual ancestral correlations will be given by

$$\rho_p = \cdot 5040\left(\frac{\beta}{\sqrt{2}}\right)^{p-1}, \quad \rho_p = \cdot 5085\left(\frac{\beta}{\sqrt{2}}\right)^{p-1}, \text{ respectively.}$$

Or, by

$$\rho_p = \cdot 5040\,(\cdot 706{,}936)^{p-1}, \quad \rho_p = \cdot 5085\,(\cdot 706{,}417)^{p-1},$$

or,

$$\rho_p = \cdot 5040\,(\cdot 080{,}318)^{p-1}, \quad \rho_p = \cdot 5085\,(\cdot 087{,}868)^{p-1}.$$

The latter values for the individual ancestral correlations show results with a ratio of degradation far too small for our experience of grandparental and great-grandparental correlations. The second roots of the quadratic—those from the minus sign—must be discarded. Or, we conclude that

$$\beta = \cdot 999{,}759 \text{ for sons, and } = \cdot 999{,}024 \text{ for daughters.}$$

But values obtained from statistical data and differing only so slightly from unity suggest most cogently that the true value of β is *unity*, or that actually our multiple regression falls into the special case dealt with on pp. 134—135. The biological significance of this result—applied to *somatic* character measurements—is very considerable. In statistical language we may state the principle as follows: The correlation of any "mid-parent" with every other mid-parent in any ancestral generation and with the offspring is the same. Symbolically $\rho_{p-q} = \alpha$, and is independent of p and q.

Now $\alpha = \sqrt{2}\,\rho_1/\{\beta(1+\epsilon)\} = \sqrt{2}\,\rho_1' = \cdot 5567$ for sons and $\cdot 5616$ for daughters.

Had there been no assortative mating ($\epsilon = 0$) we should have found

$$\alpha = \cdot 7128 \text{ for sons and } \cdot 7193 \text{ for daughters.}$$

Hence according to the intensity of assortative mating there may be considerable variation in the value of this constant mid-parental correlation.

The values obtained for mean individual ancestral correlations are as follows—of course for stature:

Offspring and	Sons	Daughters
Parent	$\cdot 5040$	$\cdot 5085$
Grandparent	$\cdot 3563$	$\cdot 3592$
Great-Grandparent	$\cdot 2519$	$\cdot 2538$
Great-great-Grandparent	$\cdot 1781$	$\cdot 1793$
Great-great-great-Grandparent	$\cdot 1259$	$\cdot 1266$
Great-great-great-great-Grandparent ...	$\cdot 0890$	$\cdot 0895$

We have no ancestral correlations for any accurately measurable character in man. Unfortunately in the case of the coat-colour in horses and eye-colour in man in the higher relationships lack of data did not permit of a separation of the sexes. The most we can do on comparison is to say that values more or less similar to the above have been found, but that until we have measurable characters in half-a-dozen generations, we can only assert that the above values for stature do not present any great improbability as to the rate of degradation of ancestral influence.

Turning to the fundamental multiple regression equation for $\beta = 1$ on p. 135, we note that for n ancestral generations:

$$\bar{x}_0 - \bar{x}_0 = \sigma_0 \frac{\cdot 5567}{1 + (n-1)\cdot 5567} \overset{n}{S}_1 \frac{X_p - \bar{X}_p}{\Sigma_p} \text{ for sons,}$$

and

$$\bar{x}_0 - \bar{x}_0 = \sigma_0' \frac{\cdot 5616}{1 + (n-1)\cdot 5616} \overset{n}{S}_1 \frac{X_p - \bar{X}_p}{\Sigma_p} \text{ for daughters.}$$

Here

$$\Sigma_p = \sigma_p \left(\frac{1}{\sqrt{2}}\right)^p \sqrt{1+\epsilon}$$

$$= 1\cdot 13155\,(\cdot 70711)^p\, \sigma_p$$

is the standard deviation of the pth mid-parents and σ_p is the standard deviation of male individuals in the pth ancestral generation. All female individuals are to be reduced to their male equivalents by multiplying their deviation from the female mean of their generation by the factor σ_p/σ_p', where σ_p' is the female standard deviation. $\bar{X}_p'$ is the mean of all possible mid-parents in the pth generation thus reduced. This definition of mid-parent and mid-parental mean is not an arbitrary one, but flows at once from the general principles of linear regression. For our particular data σ_0 is 2″·61890 for sons and σ_0' is 2″·63572 for daughters. While, as before stated, the substitution of Σ_p would appear formally to give a divergent series, this is not so actually. Every $X_p - \bar{X}_p$ must be estimated in terms of its special standard deviation, and while Σ_p gets smaller and smaller so does $X_p - \bar{X}_p$. This flows from the simple fact that without special inbreeding it is impossible to obtain several hundred ancestors in the 7th or 8th generation all deviating by the same considerable quantity from the mean of their generation. It is easy to have two exceptional parents, but not easy to have a multitude of exceptional ancestors of a high ancestral order. For particular purposes it will always be adequate to take n a simple digit in the above formulae.

To some biologists this *Note*—if they consent to read it—may appear a mere juggling with figures, which no doubt we shall be informed are valueless as data because, it will be asserted, they were not collected by a biologist! But at bottom there is a broad biological principle in the matter dealt with, which is really not beyond the comprehension of the non-mathematical. Let us take all the possible offspring of a human pair as a single group G_0, the parents as a group G_1, the grandparents as a group G_n and the ancestors in the pth generation as a group G_p. Then a study of the reduced variability of group G_0 as compared with the general population of their generation in the case of stature has shown us that if we treat

the groups G_0, G_1, ... G_n,... as individuals, these individuals are all equally associated with the ultimate offspring G_0 or with any other of these ancestral groups. It is precisely as if each group of co-generational ancestry including the offspring formed a single individual, and these single individuals like a group of brothers had the same degree of resemblance to each other. Is the individual more like his parents than his grandparents, or more like his grandparents than his great-grandparents? He is more like his *individual* parent than his *individual* grandparent, and more like the latter than his *individual* great-grandparent and so on (see p. 139). But if we drop the word *individual*, and replace it by the group as a whole, then the offspring resemble in an equal degree each successive ancestral group as represented by the appropriate "mid-parent." The individual offspring is as like to his grandparents, or indeed to his 2048 group of ancestors in the 10th preceding generation, as he is to his direct parental pair. Such is the conclusion we can draw, for at least one somatic character, by a full investigation of the multiple correlation coefficient of an indefinitely large group of brothers.

I have endeavoured in this *Note* on Mr Irwin's paper to express the ideas of which I was more or less conscious when I wrote my paragraph on the application of Galton's Difference Problem to pairs of brothers in the summer of 1902. My wording was possibly obscure, but I was convinced that the suggestion contained more than Mr Irwin is now inclined to read into it, and I am very grateful to him for bringing me back to the study of my note on the subject in the first volume of this Journal. It was not till I had read it carefully again, that my ideas of 1902 came back to me more clearly, and he has enabled me at last to put on record what I meant then by the paragraph which I venture to think he has misread.

HEIGHTS AND WEIGHTS OF PATIENTS IN MENTAL HOSPITALS.

By ROSA GREENWOOD, CECILY M. THOMPSON and HILDA M. WOODS.

Introduction.

In 1923 a Committee appointed by the Chairman of the Board of Control to consider the diets of County and Borough Mental Hospitals requested a number of Medical Superintendents to furnish particulars of heights, of weights on admission and subsequently of ages of samples of their patients.

The primary object of this request was to ascertain whether any relation could be discerned between the weights of adult patients and the computed energy values of the standard dietaries of the institutions sampled. For reasons stated in the Report of the Committee this object was not satisfactorily attained, but, as the returns provided a larger mass of statistics of heights and weights than has so far been published in reports or papers upon persons treated in English Mental Hospitals, it was thought expedient to analyse the data as completely as practicable. A note on the chief results was printed among the appendices to the Report of the Committee on Diets. That note was written by Dr Major Greenwood, a member of the Committee, under whose general supervision this analysis was made.

The inquiry has been conducted with the approval of the Committee on Quantitative Factors of Human Nutrition (Medical Research Council), and the expenses have been defrayed by a grant from the Medical Research Council.

Material.

Data of stature were furnished by seventeen mental hospitals included in the diet study and by another mental hospital not so included. Scrutiny of these data leads to the conclusion that they are reasonably accurate and provide a satisfactory basis of comparison. In the first mentioned seventeen Mental Hospitals the ages of the patients were also recorded.

Data of weights are more scanty and considerably less accurate. All the seventeen Mental Hospitals included in the diet study furnished statistics of (*a*) weight on admission, (*b*) weight at the date of making the record, but only six of these gave the weight without clothing and the estimates of the weight of clothing in the remainder were so conflicting that, save for one purpose, it was necessary to discard the statistics. The mental hospital not included in the diet study provided satisfactory records of weight, but not weight on admission.

In this paper we treat the subjects under the following headings:

1. Stature of the Mental Hospital Population Compared with Other Mental Hospital and Normal Populations.

2. Weights of Mental Hospital Patients Compared with the Weights of Normal Persons, taking account of Height and Age.

3. Changes of Weight in Inmates of Mental Hospitals During Residence.

STATURE.

Excellent material for comparison has been fully analysed by biometricians, notably Dr Tocher and the late Dr Charles Goring, whose work on the anthropometry of criminals is classical*.

In Table I*a* we set out our own results and in Table I*b* those of previous students. It will be seen on comparison of our results with those of Tocher that the agreement is good. Tocher distinguished a "General Insane" from an "Entire Insane" population, the former derived from the latter by exclusion of individuals "considered by their medical attendants to have characters affected by special causes not characteristic of insanity in general" (Tocher, p. 301)†. We have excluded no cases, so that our sample approximates more closely to the "Entire Insane" than to the "General Insane" of Tocher's grouping.

TABLE I*a*.

Mean Height with Probable Errors and Standard Deviations.

Asylum		Males			Females		
		No. of observations	Mean Height	Standard Deviation	No. of observations	Mean Height	Standard Deviation
Northern	T	182	66·27 ± ·148	2·96	183	61·79 ± ·143	2·87
	S	133	65·97 ± ·137	2·34	—	—	—
	P	249	65·19 ± ·107	2·50	231	60·64 ± ·136	3·06
	M	175	65·91 ± ·144	2·82	189	60·04 ± ·157	3·20
	F	174	65·74 ± ·200	3·90	215	61·46 ± ·148	3·22
	G	146	65·70 ± ·123	2·21	183	62·27 ± ·143	2·87
Eastern	A	217	65·52 ± ·133	2·91	230	61·06 ± ·133	3·00
	B	149	65·35 ± ·183	3·30	195	61·04 ± ·145	3·00
	N	126	65·37 ± ·179	2·98	152	61·02 ± ·130	2·38
	O	63	64·96 ± ·248	2·92	60	64·02 ± ·164	1·88
Midland	E	39	64·25 ± ·519	4·80	83	61·42 ± ·193	2·60
	D	95	64·53 ± ·222	3·20	135	59·89 ± ·158	2·73
	R	244	64·98 ± ·109	2·52	250	61·15 ± ·142	3·33
South Western	K	61	65·00 ± ·195	2·25	95	62·06 ± ·233	3·37
	H	110	65·49 ± ·182	2·83	151	60·80 ± ·185	3·37
South Eastern	J	156	65·62 ± ·181	3·35	248	62·10 ± ·126	2·95
	L	152	66·02 ± ·172	3·15	132	61·45 ± ·164	2·79
All Asylums		2471	65·55 ± ·041	3·05	2732	61·27 ± ·040	3·10

* Goring, *The English Convict*. London, 1913 (Stationery Office).

† Tocher, "Anthropometry of Scottish Insane," *Biometrika*, Vol. v. 1906, p. 298.

TABLE I*b*.

Males. Mean Stature (in inches) and its Probable Error for Professional Classes, Air Force Candidates, Convicts, Insane, and certain Occupational Classes.

	Various classes in order of mean Heights	Commercial classes, Clerks and Shopkeepers	Selected classes, Soldiers, Policemen, Messengers, Servants	Labourers: Agricultural, Roads, Quarries, Railways	Sailors, Fishermen	Miners, Coal and Mineral	Artisans	Factory Operatives
Schuster's Results (Oxford Undergraduates, *Biometrika*, Vol. VIII)	69·49 ± ·057	—	—	—	—	—	—	—
British Association Report, 1881. Population, Ages 25–30	69·14 ± ·15	67·95 ± ·12	(	67·51 ± ·05		)	66·61 ± ·09	65·92 ± ·11
Macdonell's Cambridge Students (*Biometrika*, Vol. v. p. 347)	68·86 ± ·05	—	—	—	—	—	—	—
Pearson's English Sons (*Biometrika*, Vol. v. p. 347)	68·86 ± ·05	—	—	—	—	—	—	—
Air Force Candidates (Cripps, Greenwood and Newbold, *Biometrika*, Vol. XIV. p. 316)	68·83 ± ·048	—	—	—	—	—	—	—
Air Force Candidates passed Fit (Cripps, Greenwood and Newbold, *Biometrika*, Vol. XIV. p. 316)	68·82 ± ·054	—	—	—	—	—	—	—
Roberts, General Population, Ages 23–30	68·66 ± ·12	—	67·93 ± ·03	—	—	—	66·43 ± ·12	—
" " " " 23–50	67·93 ± ·15	67·28 ± ·10	67·22 ± ·02	67·10 ± ·04	66·37 ± ·09	66·91 ± ·19	66·77 ± ·09	66·89 ± ·14
Galton's S. Kensington Lab. Data, Pearson's *Chances of Death*, Vol. I. p. 311	(67·90 ± ·06	)	—	—	—	—	—	—
Pearson's English Fathers (*Biometrika*, Vol. v. p. 347)	67·74 ± ·05	—	—	—	—	—	—	—
Tocher's General Insane Population (*Biometrika*, Vol. v. p. 303)	65·9 ± ·029	—	—	—	—	—	—	—
Tocher's Entire Insane Population (*Biometrika*, Vol. v. p. 303)	65·7 ± ·029	—	—	—	—	—	—	—
English Mental Hospitals' Data	65·55 ± ·041	—	—	—	—	—	—	—
Bexley Mental Hospital	65·54 ± ·062	—	—	—	—	—	—	—
Convicts (Goring)	65·46 ± ·036	—	—	—	—	—	—	—
Tocher's Roxburgh and Selkirk Volunteers (*Biometrika*, Vol. v. p. 347)	—	—	67·89	—	—	—	—	—
Tocher's Aberdeenshire Rural Population (*Biometrika*, Vol. v. p. 347)	—	—	—	67·72	—	—	—	—
British Recruits (*Report on Health of the Army*, 1909)	—	—	66·22 ± ·01	—	—	—	—	—
British Association Report, 1881. All males over 25	67·92	—	—	—	—	—	—	—

It appears that the males of the English Mental Hospitals are very slightly shorter on the average and the females slightly taller than those of the Scottish Mental Hospitals.

Tocher inquired whether the differences between the means of the twenty-two Mental Hospitals supplying his data were consistent with random sampling, and concluded that the inter-local differences were significant.

We have used a method of analysis similar to that of Tocher (see Tocher, *op. cit.*, p. 315 *et seq.*), and likewise conclude that there is inter-local differentiation (Table II). In the lower half of the table we set out the distribution of differences

TABLE II.

Differences of Heights from General Mean.

Asylum		Males			Females		
		Difference	Probable error of difference	Ratio of S.D. of sampling to difference	Difference	Probable error of difference	Ratio of S.D. of sampling to difference
Northern	T	+ ·72	·1525	+3·41	+ ·52	·1545	+ 2·52
	S	+ ·42	·1783	+2·09	—	—	—
	P	− ·36	·1303	−2·37	− ·63	·1375	− 3·27
	M	+ ·36	·1555	+1·74	−1·22	·1521	− 5·45
	F	+ ·19	·1559	+ ·68	+ ·20	·1426	+ ·95
	G	+ ·15	·1702	+ ·82	+1·00	·1545	+ 4·85
Eastern	A	− ·03	·1396	− ·16	− ·20	·1378	− 1·05
	B	− ·20	·1685	− ·77	− ·23	·1497	− 1·11
	N	− ·18	·1832	− ·70	− ·25	·1696	− 1·30
	O	− ·59	·2591	−1·62	+2·75	·2699	+11·23
Midland	E	−1·30	·3293	−1·71	+ ·15	·2294	+ ·53
	D	−1·02	·2110	−3·17	−1·38	·1799	− 5·89
	R	− ·57	·1317	−3·63	− ·12	·1322	− ·60
South Western	ζ	− ·55	·2633	−1·91	+ ·79	·2145	+ 2·33
	I	− ·06	·1961	− ·23	− ·47	·1701	− 1·77
South Eastern	J	+ ·07	·1646	+ ·27	+ ·83	·1327	+ 4·63
	L	+ ·47	·1688	+1·90	+ ·18	·1820	+ ·76

Males			Females		
P.E.	Observed	Theoretical	P.E.	Observed	Theoretical
±1	7	8·5	±1	4	8·0
±1—2	5	5·5	±1—2	4	5·2
±2	5	2·9	±2—3	2	2·1
			±3—	6	·7
Total	17		Total	16	
	$P = ·408$			$P = ·040$	

as multiples of the Probable Errors in comparison with the frequencies of a Normal Distribution and the deduced values of P. The number of observations is, however, very small and little importance attaches to the values of P computed from so short a series.

The range of mean statures is from just over 66·0 inches to 64·25 in males, and from 64·0 to 59·9 in females.

The mean stature of male convicts is nearly the same as that of our male patients, and their range (see Goring, p. 195) similar to that found in the sample of Mental Hospitals.

The explanation tendered by Goring to account for the inferiority of the convicted criminal population in stature, viz. that they represented an unfavourable selection of the general population, may, we think, be applied with even greater plausibility to the case of inmates of mental hospitals, a proportion of whom are known to have suffered from defects of development.

The concordance of our general results with those of Tocher is such that we have little doubt in inferring that the constants here recorded do typify the "Entire Insane" population of English Mental Hospitals, and even less doubt in concluding that this population is physically select, its males being, on the average, shorter than their social and racial class. Of course the absolute difference may be small. There is (see Table I*b*) a wide range of stature in normal populations, and it is probable—but not certain—that the population from which a majority of patients in public mental hospitals are drawn approximates in characters more closely to the Artisan Class of Roberts' survey than to any other group.

The occupational distribution of the patients forming the sample analysed was not given, but the Board of Control supplied us with the classification of the whole mental hospital population for 1908–12, which was compared with that of the General Population shown in the census of 1901. In his study of criminals, Goring employed the occupational classification of Roberts* and used the latter's standard population, viz. that of the census of 1871. We have not found in Roberts' book any definition of his categories which would enable us to group the census returns of 1901 by his method, and we think that in any case 1871 is too remote from the epoch of our observations to be a very good standard. We have therefore re-grouped Goring's data (obtained from the schedule of measurements which was separately published) and set out in Table III the three proportional distributions.

Obviously neither the Convicts nor the Asylum Patients are random samples of the General Population. It must of course be remembered that a large discrepancy under headings XXII and XXIII would be expected, having regard to the fact that neither the Convict Prisons nor the Mental Hospitals will contain a normal proportion of persons between the ages of 10 and 20.

With reference to other groups showing wide differences, both Convicts and Mental Hospital patients show larger proportions of persons engaged in Professional Occupations and in National Defence, and both show smaller proportions

* Roberts, *Manual of Anthropometry*. Churchill, London, 1878.

of miners and of textile workers. The Convicts show a defect of agricultural workers and of persons employed in the food, tobacco and drink industries; in each of these cases the Mental Hospital patients show percentages nearer to the standard. On the other hand, the proportion of Mental Hospital patients drawn from the Building Trades differs more from the standard (in defect) than does that observed among Convicts.

TABLE III.

	Occupational Group (Males)	England and Wales Population Census 1901	Asylum Population 1908–12	Criminals (Goring) Population 1902–8
	Total population less children under 10 years	12,134,259	10,304	3,000
		Rates per 1000	Rates per 1000	Rates per 1000
I	General or Local Government ...	14·1	11·4	3·3
II	Defence of Country	13·9	28·2	27·7
III	Professional	25·7	39·9	48·2
IV	Domestic	25·1	29·6	27·7
V	Commercial	43·7	57·8	71·7
VI	Conveyance of men, etc. ...	102·9	93·0	90·7
VII	Agriculture	88·3	75·1	29·7
VIII	Fishing	2·0	2·3	6·7
IX	Mines	65·9	46·5	48·7
X	Metals, Machines, etc.	96·8	63·0	73·0
XI	Precious metals, jewels, etc. ...	10·8	6·9	17·0
XII	Building	85·9	69·3	87·2
XIII	Wood, Furniture, etc.	19·2	17·1	18·3
XIV	Brick, Cement, Pottery, etc. ...	11·7	4·9	5·7
XV	Chemical, Oils	8·4	4·8	5·0
XVI	Skin, Leather, etc.	6·6	5·8	3·7
XVII	Paper, Stationery	15·5	14·2	9·0
XVIII	Textiles	40·6	24·9	12·7
XIX	Dress	34·2	39·9	53·0
XX	Food, Tobacco, Drink	63·8	53·5	37·0
XXI	Gas, Water, etc.	5·9	4·4	3·3
XXII	Other general and undefined ...	56·0	200·2	285·7
XXIII	Without specified occupation or unoccupied	163·0	107·3	35·0

A perusal of this table does not suggest that the occupational groups manned by the middle and upper classes are under-represented in the Mental Hospital Population, and therefore does not offer an obvious explanation of any inferiority of physique observed. Thus the same conclusion as drawn by Goring (*op. cit.*, p. 178) holds, but less securely than in his case because we have not been able, as he was, to compare mental hospital patients and normal persons from the same occupational class. This is a task for the future, one thoroughly worth accomplishment.

In Table IV we give the correlation coefficients between stature and age, and the correlation ratios of stature upon age for normal and mental hospital populations. The differences are not such as to merit comment.

TABLE IV.

Correlation Coefficients and Ratios of Age and Height for Air Force and Insane.

	No. of Observations	Mean Age (Yrs.)	Standard Deviations	Mean Height (inches)	Standard Deviations	Correlation Coefficients	Correlation Ratios
Air Force Candidates. Males.	1238	23·92 ± ·097	5·07	68·83 ± ·048	2·50	− ·0339 ± ·019	·1790 ± ·019
English Mental Hospital. Males.	2471	41·94 ± ·209	15·41	65·55 ± ·041	3·05	·0092 ± ·014	·2298 ± ·013
English Mental Hospital. Females.	2732	43·68 ± ·199	15·42	61·27 ± ·040	3·10	− ·0777 ± ·013	·2654 ± ·012

Body Weight.

The subject of change of body weight with age and stature is evidently one of considerable importance. Its importance in *individual* prognosis is, of course, a commonplace of medical experience, since a deviation from the individual's norm affords valuable and early indications of disease. The value of a comparison of the weight of an individual with the standard provided by statistical averages of persons of his age, stature, race and social class is undoubtedly smaller, but the very extensive investigation published by the Medico-Actuarial Society of America* shows that even here importance can be attached to the procedure. It therefore seemed that results of interest would flow from an analysis of experience in Mental Hospitals. Unfortunately we have been hampered by two serious defects. The first is that data, even for normal populations, are less adequate than might have been expected; the second, that the measurements taken in the Mental Hospitals have been, on the whole, unsatisfactory.

The literature of adult body weight is far more scanty and much less exact than that of stature; the weight of adult women has indeed hardly been discussed at all in England and Wales during recent years. Most of the statistics of body weights relate to *clothed* persons, and the weights of the clothing have often been estimated in an unsatisfactory way. In the case of the Mental Hospitals, the estimates of weight of clothing had a range from 7 to 14 lbs., and we were quite unable to accept these estimates as having any pretensions to exactitude.

The best data regarding body weight in healthy adults to which we had access were measurements of 1238 Air Force candidates (Cripps, Greenwood and Newbold)†; these men were weighed stripped, and had a mean weight of 139·86 lbs. Goring's weights of 2500 convicts, weighed in trousers and shirt, are probably equally accurate. Their mean weight was 142·56 lbs., but their mean age was

* Medico-Actuarial Society of America. *Mortality Investigation. Statistics of Height and Weight of insured Persons.* New York, 1912, Vol. I.

† Cripps, Greenwood and Newbold, "A Biometric Study of the inter-relations of 'Vital Capacity,' Stature, Stem length and weight in a sample of Healthy Male Adults," *Biometrika*, Vol. XIV. Parts 3 and 4, 1923, p. 316.

36·24 years and the mean age of the Air Force candidates was only 23·92 years, while the statures differed by more than 3 inches. In general there is a positive correlation between weight and age, and between weight and height. Since the regression of weight upon age is not linear, the customary linear equations of regression are not exact methods of graduation, but they may be used to make rough allowance. Using the regression equation computed on the Air Force data, it appears that an Air Force candidate of the mean age and stature of the convicts would weigh 137 lbs. Allowing 2 lbs. (a minimum estimate) for the weight of trousers and shirt, we reach as the figure to be compared with Goring's, 139 lbs.

All other data relate to clothed men. The 3000 measurements made by Hutchinson* eighty years ago had a mean of 147·86 lbs. The 5032 men, aged 25 years or more, reported on by the British Association† forty years ago, had an average of 159·2 lbs. Lastly, Schuster's‡ weighings of undergraduates had a mean of 151·94 lbs. (these men were weighed without boots or coat).

Hutchinson's subjects included persons between 15 and 40, all but fifty-nine of Schuster's subjects were between 18 and 23, so that both samples might be expected to give lower averages than that of the British Association, which covered a larger proportion of older persons.

The Medico-Actuarial Society of America data refer to 221,694 men having a mean age of 33 years and a mean stature of 68·5 inches. The average weight—that of the clothed man—was 156·2 lbs. Levine§ has recently communicated the results of 28,697 examinations for life assurance in Great Britain and Ireland during the years 1921–22 (these, of course, are also weights of clothed men). From the graduated results, it appears that for a man of 36 standing 65½ inches (the mean values for Goring's convicts), the mean weight would be about 142 lbs. This is very near Goring's figure, and perhaps, since the candidates for assurance probably had more than shirt and trousers on, is evidence of a slightly lower mean weight.

Owing to the difficulty about clothing, we are really not in a position to state precisely what the average weight of a normal adult is. But, making allowance for stature, we think that the Air Force candidates were somewhat below the general average or, alternatively, that predictions based upon the Air Force data—involving the assumption of linear regression—slightly under-state the position, if the British Association data are taken as a standard. Thus the average age of the males over 25 in the British Association's experience was 34·9 years and the average stature 67·92 inches. Using the equation (deduced from the Air Force records)

$$\text{Weight (in pounds)} = 0{\cdot}7326\ \text{Age (in years)} + 3{\cdot}6133\ \text{Height (in inches)} - 126{\cdot}37,$$

* Hutchinson, "On the capacity of the lungs and on the Respiratory Functions with a view to establishing a precise and easy method of detecting disease by the Spirometer," *Transactions of the Royal Medico-Chirurgical Society of London*, Vol. XXIX.

† *British Association, Report* 1881. Report of the Anthropometric Committee, pp. 242–3.

‡ Schuster, "First Results from the Oxford Anthropometric Laboratory," *Biometrika*, Vol. VIII. 1911, p. 40.

§ Levine, "Heights and Weights," *Journal of the Institute of Actuaries*, Vol. LIV. Part 2, July 1923, p. 213.

we predict for this age and height, 144·6 lbs., or 14·6 lbs. below the observed weight. 14·6 lbs. seems an improbable figure for the weight of clothing. But on the Assurance Standard, the agreement is close. Levine's graduation would put the weight for this age and height at about 151 lbs., which only requires somewhat more than 6 lbs. to be allowed for clothing to coincide with the Air Force prediction. We do not of course know whether the older British Association figures or the modern assurance figures more correctly typify the normal average. If we accept the latter, there is very good concordance between the Air Force data and the assurance data, and, making allowance for difference of age and stature, between both and the weighings of criminals.

Upon the whole, therefore, we seem quite justified in judging our mental hospital data for males with the help of the equation deduced from the Air Force records.

Data of weights of women are, for England and Wales, extremely scanty. In Table V we give the results of the American Medico-Actuarial Investigation

TABLE V.

Means and Standard Deviations of Age, Weight, and Height for various groups of Women.

	Age (in years)			Weight (in lbs.)			Height (in inches)		
	Observations	Means	Standard Deviations	Observations	Means	Standard Deviations	Observations	Means	Standard Deviations
American Medico-Actuarial Investigation	136375	32·571	9·149	136375	134·196	19·428	136375	64·282	2·300
Cripps — Polytechnic Physical Training School ...	232	20·52	2·29	226	124·38	11·95	226	64·78	1·58
Cripps — Somerset House ...	124	26·34	7·69	125	117·24	16·91	125	62·96	2·32
Cripps — Bedford College ...	115	21·13	1·79	70	117·75	16·29	70	64·42	2·52

Equation for Polytechnic Scholars: Weight = 3·9072 height − 128·73.
Equation for American Women: Weight = 3·4978 height + ·6437 age − 111·6156.

and three short series of English data. The difficulty about the allowance to be made for clothing is so great that any comparison is hazardous, but the following considerations suggest that the average weight of American women, allowance being made for height and age, is greater than that of English women. From the American data we have calculated the regression equation connecting weight, height and age and have applied this to the prediction of weights for the given mean ages and heights in three English series. We find for the Polytechnic women 128·1 lbs. compared with 124·4 observed, for the Somerset House women 125·6 against 117·2 observed, and for the Bedford College students 127·3 as against 117·8 observed. The Polytechnic Physical Training School women were as a whole particularly vigorous subjects and, bearing in mind both the difficulty about clothing and the fact that the range of age covered was very small, we cannot

TABLE VI.

Comparison of Weights on Admission and at Time of Inquiry in Six Asylums where weights were made without clothes.

Asylum		No. of Observations	1st Weight (lbs.)		2nd Weight (lbs.)		Difference 1st & 2nd weight	Percentage who			Length of residence (months)	Age (years)	Calories (from dietaries)	Correlations of Height and Weight
			Mean	Standard Deviation	Mean	Standard Deviation		gained	lost	no change				
Northern	G M	146	133·20±1·06	19·05	138·81±1·23	22·11	5·62	70	26	4	12·21	41·73	2926	+·5741±·037
	G F	183	115·52±1·05	21·09	119·62±1·09	21·92	4·10	58	30	12	10·72	47·28	2582	+·3730±·043
	P M	249	132·00± ·72	16·89	140·67± ·82	19·15	8·68	74	24	2	18·58	40·70	2249	+·4839±·033
	P F	231	107·46± ·96	21·55	115·38±1·00	20·47	7·92	72	26	2	16·22	39·59	2006	+·2544±·042
	T M	182	121·05± ·92	18·40	125·47± ·94	18·79	4·42	62	34	4	12·83	43·03	2553	+·6173±·031
	T F	183	104·92± ·87	17·41	108·04± ·97	19·38	3·12	67	29	4	13·90	41·67	2218	+·2994±·045
Eastern	N M	126	133·27±1·33	22·07	136·12±1·33	22·10	2·86	59	39	2	12·02	43·20	3144	+·5119±·044
	N F	152	112·29±1·27	23·21	112·26±1·23	22·50	−·03	53	45	1	12·39	44·25	2796	+·1994±·053
Midland	D M	95	118·62±1·66	24·05	130·41±1·62	23·39	11·79	83	16	1	12·20	43·24	2239	+·6251±·042
	D F	135	107·32±1·15	19·87	109·99±1·14	19·70	2·67	60	36	4	11·13	43·98	1902	+·4152±·048
	R M	244	124·55± ·83	19·29	132·79± ·82	19·09	8·24	76	21	3	13·68	38·82	2844	+·4630±·034
	R F	250	109·77± ·92	21·53	118·37± ·94	22·06	8·60	70	27	3	17·28	41·50	2565	+·3206±·038
Six Asylums (Weighted Averages)	M	1042	127·44± ·40	19·29	134·43± ·44	20·95	6·98	71	26	3	14·16	41·34	2644	+·4991±·016
	F	1134	109·49± ·42	20·83	114·48± ·43	21·47	4·99	65	31	4	14·07	42·74	2350	+·3099±·018

be confident that their mean weight is really much less than that of the Americans. But the difference in both the other groups is so large that it seems most improbable that the true mean of either reaches the American Standard.

Weights of Mental Hospital Patients.

In Table VI are given the particulars of weights for both males and females in six hospitals which supplied unclothed weights.

In Table VII we compare the weights of the males with the predictions from the Air Force equation, allowing for age and height.

TABLE VII.

Males. Means of Final Weights in Table IV compared with estimates from Air Force Equation.

Asylum		Observed weight	Predicted weight	Difference
Northern	G	138·8 ± 1·23	141·6 ± ·33	+ 2·8
	P	140·7 ± ·82	139·0 ± ·26	− 1·7
	T	125·5 ± ·94	144·7 ± ·30	+19·2
Eastern	N	136·1 ± 1·33	141·6 ± ·36	+ 5·5
Midland	D	130·4 ± 1·62	138·3 ± ·88	+ 7·8
	R	132·8 ± ·82	136·9 ± ·26	+ 4·1
Whole Sample		134·4 ± ·42	140·7 ± ·13	+ 6·3

The equation: Weight (lbs.) = ·7326 age (years) + 3·6133 height (inches) − 126·37.

In one instance, the observed weight exceeds the predicted weight, in the other five it is in defect.

Confining ourselves to persons over 20 years, the six mental hospitals provided 969 males and 1084 females. The biometric constants were evaluated and it appeared that the correlation of age with height was negligible. In Table VIII we show the mean weights for arrays of age together with the values assigned by the regression of weight upon age.

When these are tested in the usual way for Goodness of Fit* it appears—as indeed is obvious from the table—that the fit (in males) is bad, but that this is due to great divergence in particular groups. Thus, in the case of females, 38·6 % of the value of χ^2 is contributed by the last group, and its omission would increase P from ·25 to ·65 ($n' = 14$). In males the last two groups contributed nearly 62 % of the value of χ^2 and their omission improves P to ·073 (for $n' = 13$), and even in males the general trend is fairly well represented by a straight line. In other words, although the regression is certainly not linear, we may say that a comparison of the regression coefficients of weight upon age between

* *Biometrika*, Vol. IX. 1913, p. 28.

equations derived from normal populations and from mental hospital populations is a justifiable proceeding. Such comparison shows us that the increase of weight

TABLE VIII.

Weights at different Ages for 969 *Males and* 1084 *Females in Six Asylums.*

Years of age	Weights of Males (lbs.)			Weights of Females (lbs.)		
	No. of Observations	Observed	Calculated	No. of Observations	Observed	Calculated
23	82	131·23	131·62	99	113·13	113·37
28	129	131·21	132·58	116	112·02	113·65
33	138	137·35	133·54	135	112·06	113·94
38	125	130·93	134·50	133	114 87	114·23
43	106	138·55	135·46	150	114·32	114·51
48	112	138·29	136·42	128	117·05	114·80
53	76	134·33	137·38	87	118·58	115·09
58	70	141·54	138·33	77	117·59	115·37
63	61	138·20	139·29	71	113·42	115·66
68	39	136·28	140·25	39	112·69	115·95
73	23	137·86	141·21	26	116·40	116·23
78	5	147·25	142·17	20	110·25	116·52
83	2	105·25	143·13	2	110·25	116·8
88	1	95·25	144·09	1	85·25	117·09

Equations: Weight = ·1917 age + 127·2156 Weight = ·0573 age + 112·0487

$n' = 15$ $n' = 15$

$x^2 = 50{\cdot}3658$ $x^2 = 17{\cdot}0964$

$P =$ ·00 $P =$ ·25

with age is much smaller amongst the insane than amongst normal persons. While in respect of stature the insane are differentiated from the normal to the same extent as convicted criminals, the latter are not, in respect of weight, subnormal, the former are. The most important cause of this is, undoubtedly, that the population of a mental hospital contains many more physically diseased persons than a convict prison. Goring's analysis of the vital statistics of criminals showed *inter alia*, that "the current allegation of criminality and tubercular disease being kindred manifestations of the same form of human decadence is entirely unsupported by the statistical facts" (Goring, *op. cit.*, p. 229).

In mental hospitals, on the other hand, the rate of mortality and the prevalence of tuberculosis are very high. We cannot indeed account for our low average weights by the presence in the samples of overtly tuberculous patients, the numbers so recorded were very few, but that many were in a pre-tuberculous state can be regarded as probable. This, of course, at once raises the extremely interesting question, which is cause and which effect, but the present data are neither numerous enough nor accurate enough to allow us to discuss that question; a further peculiarity of the material is dealt with below. Our demonstration of the low weight for age and stature of the population is, we submit, a cogent argument in favour of requiring that changes of body weight in mental hospitals should be made the object of much more exact inquiry.

We are not acquainted with any published investigations of changes of weight in the chief clinical types of insanity and, having regard to the enormous field for investigation offered by the mental hospitals of this country, lack of such information cannot be justified. It appears that contempt for very humble methods of research, mere weighings and measurings, has been carried too far. Perhaps careful anthropometry would suggest to the clinician something of as much practical importance as is to be gathered from much recent psychological speculation.

TABLE IX.

Asylum			No. of Observations	Mean Age (years)	Mean Period of observation (months)	Mean Gain in weight (lbs.)	Percentage of patients observed who lost weight while under observation
Northern	F	M	174	42·1	11·8	6·54	20·7
		F	215	46·3	11·2	2·30	38·6
	G	M	146	41·7	12·2	3·66	26·0
		F	183	47·3	10·7	4·10	30·1
	M	M	175	43·5	13·9	4·27	28·0
		F	189	41·7	16·3	6·50	26·5
	P	M	249	40·7	18·6	8·21	25·7
		F	231	39·6	16·2	7·40	25·5
	S	M	133	43·4	14·5	4·31	31·6
		F	—	—	—	—	—
	T	M	182	43·0	12·8	3·97	33·5
		F	183	41·7	13·9	3·60	29·0
Eastern	A	M	217	41·7	12·3	5·49	30·4
		F	230	42·9	15·1	9·10	22·6
	B	M	149	42·1	13·7	7·28	22·1
		F	195	43·2	13·9	8·90	24·1
	N	M	126	43·2	12·0	3·39	38·9
		F	152	44·3	12·4	− ·20	45·4
	O	M	63	47·0	19·3	2·48	33·3
		F	60	45·3	13·9	4·90	28·3
Midland	D	M	95	43·2	12·2	11·15	15·8
		F	135	44·0	11·1	2·70	36·3
	E	M	39	48·2	13·3	11·00	17·9
		F	83	48·5	12·3	6·40	26·5
	R	M	244	38·8	13·7	7·95	21·3
		F	250	41·5	17·3	5·40	27·2
South Western	H	M	110	42·3	12·2	4·23	15·9
		F	151	45·8	10·5	5·40	18·5
	K	M	61	42·5	10·2	4·52	27·9
		F	95	44·9	11·7	7·00	24·2
South Eastern	J	M	156	41·4	11·6	3·27	34·6
		F	248	44·8	13·9	− ·60	46·8
	L	M	152	40·4	11·7	3·59	36·2
		F	132	44·3	13·1	6·10	27·3
Weighted Averages		M	2471	42·0	13·4	5·60	27·2
		F	2732	43·7	13·7	4·90	30·3

THE CHANGE OF WEIGHT OF PATIENTS BETWEEN ADMISSION AND DATE OF OBSERVATION.

In the above sections we have used the weights at the date of completing the records, and these were at various times after admission.

The subject of the change in weight of patients during residence in an asylum is a matter of considerable interest. In Tables IX and X we show the difference

TABLE X.

The Changes in Weight of Patients between Date of Admission and Date of Inquiry.

Mental Hospital		Total no. of observations in each asylum	Gained Weight				Lost Weight			
			No. who gained	Mean first weight	Mean gain	Mean final weight	No. who lost	Mean first weight	Mean loss	Mean final weight
Northern	f M	174	130	122·9	11·28	134·2	36	132·0	8·28	123·7
	F F	215	120	100·3	12·39	112·7	82	120·1	10·88	109·2
	g M	146	102	132·0	11·37	143·4	38	138·9	7·23	131·7
	G F	183	107	111·7	11·12	122·8	55	124·8	7·27	117·5
	m M	175	112	131·9	11·21	143·1	50	145·6	11·46	134·1
	M F	189	134	107·3	13·19	120·5	50	118·3	9·50	108·8
	p M	249	185	129·7	14·36	144·1	59	140·9	8·97	131·9
	P F	231	166	103·5	15·09	118·6	59	121·6	11·95	109·7
	s M	133	82	130·2	9·18	139·4	43	132·5	5·07	127·4
	S F	—	—	—	—	—	—	—	—	—
	t M	182	112	117·3	10·80	128·1	61	128·5	6·59	121·9
	T F	183	122	104·0	8·37	112·4	53	107·2	7·21	100·0
Eastern	a M	217	140	121·4	11·80	133·2	66	133·1	6·28	126·8
	A F	230	173	105·3	15·29	120·6	52	120·9	8·43	112·5
	b M	149	109	125·1	12·89	138·0	33	138·1	8·09	130·0
	B F	195	146	104·8	15·47	120·3	47	118 6	11·59	107·0
	n M	126	74	128·2	12·74	140·9	49	141·4	10·89	130·5
	N F	152	81	110·4	11·03	121·4	69	115·6	13·71	101·9
	o M	63	37	128·8	10·22	139·0	21	142·6	8·19	134·4
	O F	60	40	113·3	10·65	124·0	18	120·2	7·39	112·8
Midland	d M	95	79	116·5	15·38	131·9	15	134·3	8·87	125·4
	D F	135	81	102·3	13·14	115·4	49	114·6	13·12	101·5
	e M	39	32	118·1	16·73	134·8	7	150·7	9·00	141·7
	E F	83	59	112·9	12.89	125·8	22	129·9	9·68	120·2
	r M	244	185	121·5	13·00	134·5	52	136·1	7·87	128·2
	R F	250	175	106·3	15·69	122·0	67	120·5	9·23	111·3
South Western	h M	110	74	127·5	11·11	138·6	29	133·4	10·69	122·7
	H F	151	122	109·4	9·87	119·3	28	124·0	16·63	107·4
	k M	61	41	133·3	10·48	143·8	17	143·9	12·29	131·6
	K F	95	66	111·5	14·61	126·1	23	120·4	12·25	108·2
South Eastern	j M	156	95	131·3	11·16	142·5	54	142·7	10·08	132·6
	J F	248	117	108·6	10·54	119·1	115	123·8	10·93	112·9
	l M	152	90	134·5	10·99	145·5	56	145·4	7·64	137·8
	L F	132	92	105·6	12·59	118·2	35	110·6	8·37	102·2

between the weight on admission and the final weight in all the institutions making returns, including of course the six providing what we think are accurate weights. The others were included here because although, owing to the difficulty about the weight of clothing, the absolute values are useless, it is probable that the method of weighing in any one institution is constant so that the differences between initial and final weights may have some value. Even this, however, is open to question, because some of the weighings were in the summer and some in the winter, so that differences between summer and winter clothing are possibly of importance. There is, however, no uniformity in the seasonal distribution of initial and final weighings, so that such an error is randomly distributed.

It is very important to notice that the patients weighed are a peculiar selection of the inmates. The superintendents were asked (on May 30th, 1922) to supply the information* "concerning the first 250 male and 250 female patients admitted since January 1st, 1920, in order of their admission, who are still resident in your Mental Hospital. Should the number of admissions since the date mentioned be less than 250, the actual admissions will meet requirements." Consequently—the average time of residence of the patients measured being 13 months—there are excluded from the sample (*a*) persons who died within the average period of months from admission, (*b*) persons who were discharged cured within that period. We are therefore wholly dealing with persons suffering from some mental disorder of a relatively chronic type not accompanied at admission by bodily disease of a rapidly fatal type.

For the reasons stated we can, so far as the majority of the institutions are concerned, merely express the following *prima facie* deductions from Table X.

(1) There is no relation between scheduled diet and body weight.

(2) About 30 per cent. of the sampled population lose weight between admission and the date of final weighing, which is from 12 to 18 months after admission. The mean increases and decreases shown in Table X are in several instances large, but, bearing in mind the very serious criticism to which the methods of weighing are subject, no stress is put upon the precise figures; they do but indicate the need of a wider *and more accurate* survey.

(3) This result cannot be explained by the existence in the mental hospital population, as sampled, of a supra-normal proportion of persons suffering from bodily disease—such as phthisis—notoriously associated with loss of weight. Giving a very liberal interpretation to the terms physical disease likely to lead to loss of weight, we find that the proportion of patients in our sample belonging to this class is less than 5 per cent., and simple calculation shows that dilution of normal material with 5 per cent. of persons even enormously underweight would not bring down the general average as low as it is found to be.

(4) A plausible explanation is that many victims of acute psychoses are known to lose weight rapidly and that this sample will include a relatively large propor-

* Extract of official letter of Board of Control.

tion of such cases. On the other hand, the circumstances of persons of unsound mind coming from the economically less prosperous classes are often very unfavourable, and such persons would tend to put on weight when cared for in a public institution. These two considerations tend in opposite directions, qualitatively; what their respective values are we cannot tell.

For males of the average height of our asylum patients, viz. $65\frac{1}{2}$ inches, the increase of weight for each quinquennium of age ranged from 3 to 5 lbs. between 45 and 20 and was rather less than 2 lbs. at older ages. We have seen that in the Air Force data the approximate increase for each year of age—stature constant—was rather less than $\frac{3}{4}$ lb. or somewhat more than $3\frac{1}{2}$ lbs. for a quinquennium. This is a finding in good agreement with the Insurance Offices data.

Since the correlation of age and stature (adults) is very slight, the position is sufficiently well brought out by the accompanying Diagram 1 which shows the means of arrays of weights for quinquennial ages.

Weights at different Ages for Males and Females in six good Asylums.

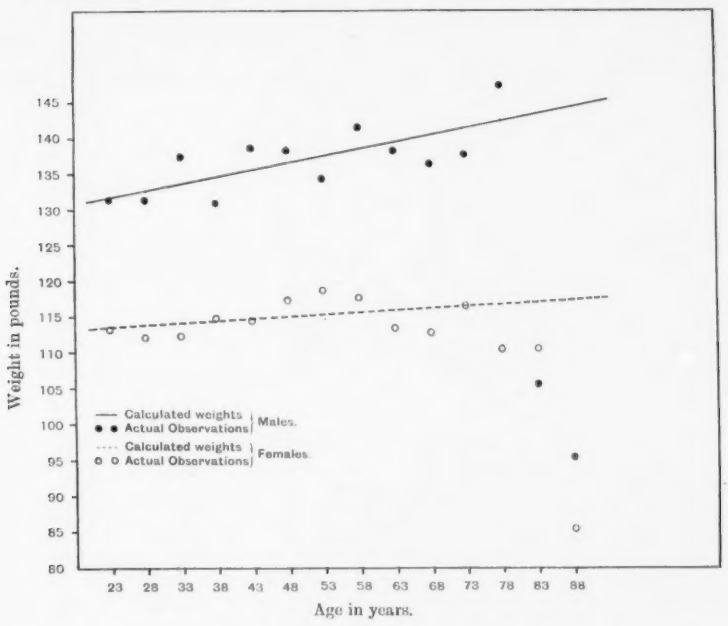


Diagram 1

Whether this divergence from the normal experience has a definite biological meaning or whether it is due to the unsatisfactory character of the material we are unable to say.

We think, however, that the results of this inquiry are sufficient to emphasize the need of a much more searching investigation of changes in weight of patients in mental hospitals than the present material has permitted us to make.

It is possible that the simplicity of the measurements has caused them to be despised as objects of scientific research, and certainly no one would claim for them a high rank in the appraisement of clinical conditions in individual prognosis or treatment. But it would, we think, be of some interest to ascertain whether the anthropometric characters of different clinical groups within the population of mental hospitals and the variations of such characters with age are significantly differentiated.

MISCELLANEA.

Some further Notes on Cancer and Goitre Distributions.

By PERCY STOCKS, M.D.

In a previous paper* evidence was collected which indicated a relation between the frequency of goitre and the incidence of cancer of the digestive organs, more especially the stomach. Since the publication of that paper some additional evidence of the same kind relating to the geographical distribution of these diseases in America and Sweden has been obtained, and is here set down.

I. *More Complete Analysis of American Statistics.*

During the Great War some 2½ millions of recruits called up for service in the U.S. Army were subjected to a careful medical examination of which the records have been analysed in an official volume†. From these statistics the number of men per thousand examined from each State who exhibited various diseases can be ascertained. As regards anomalies of the thyroid gland the disorders included are "goitre (simple)," "exophthalmic goitre" and "cretinism." The rates for the last are too small to be reliable for the purpose in view.

The distribution of exophthalmic goitre is found to be highly correlated with that of simple goitre, the correlation coefficients between the frequencies in the various States being :

For the whole area	$r = \cdot 674 \pm \cdot 054$
For 37 States east of Long. 104° W. ...	$r = \cdot 830 \pm \cdot 031$
For 35 Registration States	$r = \cdot 749 \pm \cdot 043$

This high positive correlation is difficult to reconcile with many accepted views as to the causation of Graves' disease, and is in contrast with the infrequency of this disease in goitrous areas of Switzerland. The question arises whether the dividing line between simple and exophthalmic goitres was the same in all the States, or whether in goitrous areas the Local Boards may not have looked more carefully for slight indications of hyperthyroidism and therefore classed a larger percentage of all goitres as exophthalmic than was the case elsewhere. To test this point we may separate the rejected men and those examined by Local Boards from the accepted men examined by Medical Boards in camps where the effect of any such local diagnostic peculiarity would be less: the correlations are, for rejected men $r = \cdot 611 \pm \cdot 062$, and for accepted men $r = \cdot 689 \pm \cdot 052$, from which it appears that the local factor as regards diagnosis is not of any importance. The correlation between Graves' disease and simple goitre distributions in the U.S.A. must therefore be accepted as a rather surprising fact, and it rather suggests that two distinct factors may be necessary to produce Graves' disease, viz. (i) a goitre-producing factor whose intensity is indicated by the prevalence of simple goitre, (ii) another factor, probably of nervous origin, which is only present to a slight extent in Switzerland, but is more pronounced in the U.S.A. and almost universally present in England and Wales. This, however, is mere speculation.

I have indicated in Table I and on Map No. 1 the total rate for simple goitre in each of the American States (i.e. number of men per 1000 examined by Medical Boards, whether rejected or accepted for service, in whom simple goitre was noted). In the Report of the 1900 Census the

* "Cancer and Goitre," *Biometrika*, Vol. XVI. Parts 3–4, Dec. 1924, pp. 364–401.

† "Defects found in Drafted Men." War Dept. Statistical Information, Washington, 1920.

TABLE I.

Goitre and Cancer Distribution in U.S.A.

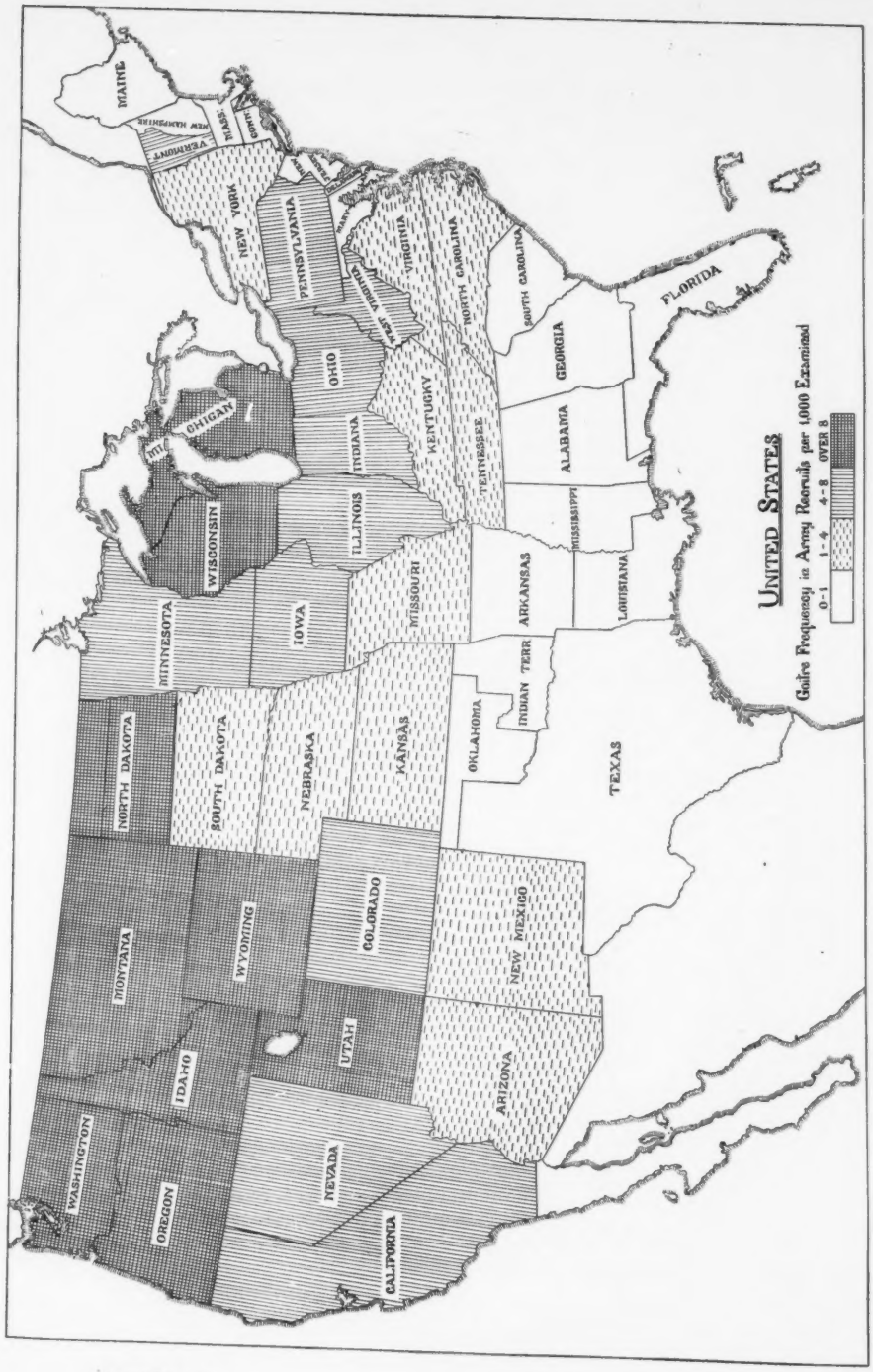
State	Simple Goitre per 1000 males (1)	Exophthalmic Goitre per 1000 males (1)	1900					1920				
			No. of deaths			Selection rate of Cancer for stomach (2)	Percent. males in population	Population	Percent. males in population	Percent. over 45	Corrected mortality rate	
			Cancer total	Cancer of stomach	Cancer, organ not stated						Cancer total (3)	Cancer stomach and liver (4)
Alabama	·56	·79	340	48	89	19·1	50·1	2,348,174	—	—	—	—
Arizona and New Mexico	1·07	·91	43	9	10	27·3	55·3	694,512	—	—	—	—
Arkansas	·40	·35	171	31	63	28·7	51·5	1,752,204	—	—	—	—
California	4·45	1·95	1026	295	271	37·7	55·2	3,426,861	52·9	26·1	92·2	42·7
Colorado	5·29	3·60	184	56	58	44·4	54·7	939,629	52·4	21·9	73·5	37·9
Connecticut ...	·89	1·53	555	95	158	23·9	50·0	1,380,631	50·4	22·2	96·1	44·6
Delaware	·59	·79	75	9	42	27·3	51·0	223,003	51·0	23·7	79·4	41·2
District of Columbia	1·39	3·47	—	—	—	—	—	437,571	46·5	22·8	85·9	40·5
Florida	·25	·66	83	19	38	42·2	52·1	968,470	51·1	19·4	55·8	23·9
Georgia	·52	1·61	373	30	141	12·9	49·8	2,895,832	—	—	—	—
Idaho	26·91	3·20	28	9	5	39·1	57·7	431,866	—	—	—	—
Illinois	7·79	5·57	2102	676	542	43·3	51·3	6,485,280	50·9	21·6	89·5	44·7
Indiana	6·49	3·12	1036	214	441	36·0	49·9	2,930,390	50·8	24·8	77·8	37·8
Iowa	6·68	2·86	801	273	253	49·8	51·8	2,404,021	—	—	—	—
Kansas	1·25	2·76	522	143	199	44·2	52·3	1,769,257	51·4	22·4	68·2	32·9
Kentucky	1·41	1·38	548	126	205	36·4	50·8	2,416,630	50·8	19·9	58·6	27·1
Louisiana	·62	·75	316	64	90	28·3	50·3	1,789,509	50·5	16·1	60·9	22·9
Maine	·66	·86	581	158	125	34·6	50·5	768,014	50·6	28·1	95·0	43·4
Maryland	·94	3·57	548	164	125	38·8	56·7	1,449,661	50·3	22·3	89·2	42·4
Massachusetts ...	·32	·63	1864	347	738	30·8	48·7	3,852,356	49·1	24·2	101·6	40·5
Michigan	11·43	6·89	1404	411	338	38·6	51·6	3,668,412	52·6	21·7	82·0	42·0
Minnesota	8·04	3·38	659	251	164	50·7	53·2	2,387,125	53·4	20·7	97·2	54·0
Mississippi ...	·64	·80	258	39	104	25·3	50·4	1,790,618	50·1	16·3	46·5	17·4
Missouri	3·99	4·80	908	229	251	34·9	51·4	3,404,055	50·6	23·1	71·1	35·2
Montana	21·00	3·39	55	22	7	45·8	61·6	548,889	54·6	18·4	57·5	27·2
Nebraska	2·14	1·70	260	82	144	45·8	52·9	1,296,372	51·9	20·5	78·2	43·1
Nevada	6·38	1·52	14	4	4	40·0	60·5	77,407	—	—	—	—
New Hampshire ...	·70	·94	275	46	156	38·7	49·9	443,083	50·1	28·5	90·0	39·1
New Jersey ...	·43	1·54	949	260	243	36·8	49·9	3,155,900	50·4	21·1	86·2	39·5
New York	1·19	2·94	4448	1097	1176	33·5	49·7	10,385,227	49·7	22·5	94·8	42·8
North Carolina ...	1·81	1·91	350	47	124	20·8	49·5	2,559,123	50·4	16·1	48·7	20·7
North Dakota ...	8·73	3·41	50	15	19	48·4	55·6	646,872	—	—	—	—
Ohio	5·59	5·08	1943	490	745	40·9	50·6	5,759,394	51·3	23·2	82·0	40·0
Oklahoma	·72	·87	49	14	12	37·8	53·8	2,028,283	—	—	—	—
Oregon	26·31	6·37	143	52	36	48·6	56·3	783,389	53·1	24·4	82·5	39·4
Pennsylvania ...	4·10	4·81	2607	688	746	36·9	50·9	8,720,017	50·8	20·9	83·5	41·9
Rhode Island ...	·55	1·24	267	57	44	25·5	49·1	604,397	49·2	23·3	90·7	37·1
South Carolina ...	·94	1·73	231	18	81	12·0	49·6	1,683,724	49·8	14·4	44·2	18·1
South Dakota ...	4·09	3·18	77	24	25	46·1	53·8	636,547	—	—	—	—
Tennessee	1·96	2·07	447	66	161	23·1	50·5	2,337,885	50·2	18·6	51·1	17·8
Texas	·30	·35	478	108	165	34·5	51·8	4,663,228	—	—	—	—
Utah	17·52	3·82	68	20	27	48·8	51·2	449,396	51·6	16·6	66·9	39·0
Vermont	2·14	4·28	281	60	115	36·1	50·9	352,428	50·8	28·5	96·4	43·9
Virginia	3·38	3·99	461	91	148	28·7	49·9	2,309,187	50·6	18·5	61·8	25·9
Washington ...	23·40	9·42	185	47	67	39·8	58·7	1,356,621	54·2	22·7	77·8	40·5
West Virginia ...	7·89	4·52	188	54	50	39·1	52·1	1,463,701	—	—	—	—
Wisconsin	14·02	7·94	980	313	390	53·1	51·6	2,632,067	51·5	22·1	85·0	50·1
Wyoming	15·38	2·71	16	6	5	54·5	62·9	194,402	—	—	—	—

(1) Total rate from all recruits examined, whether accepted or rejected.

(2) Corrected by omitting from total cancer deaths those where organ was unspecified.

(3) Corrected for age and sex distribution by approximate standardization from data in preceding columns.

(4) Corrected for cancers of unspecified site, and for age and sex by same methods as (3).


UNITED STATES
Goître Frequency in Army Recruits per 1,000 Examined
0-1 1-4 4-8 OVER 8
MAINE
VERMONT
NEW HAMPSHIRE
MASS.
NEW YORK
PENNSYLVANIA
OHIO
MICHIGAN
WISCONSIN
ILLINOIS
INDIANA
KENTUCKY
VIRGINIA
NORTH CAROLINA
SOUTH CAROLINA
GEORGIA
FLORIDA
ALABAMA
MISSISSIPPI
LOUISIANA
TENNESSEE
ARKANSAS
MISSOURI
IOWA
MINNESOTA
NORTH DAKOTA
SOUTH DAKOTA
NEBRASKA
KANSAS
OKLAHOMA
INDIAN TERR
TEXAS
MONTANA
WYOMING
COLORADO
NEW MEXICO
IDAHO
UTAH
ARIZONA
WASHINGTON
OREGON
NEVADA
CALIFORNIA

MAP No. 1.

number of deaths during that year from cancer of all organs and for cancer of the stomach were recorded for all the States, but in subsequent years mortality figures were confined to a limited number of "Registration States." In the 1920 report these number 35, and in these more recent records, cancer of the stomach and liver are grouped together. In 1900 the statistics for many of the sparsely populated Western States were of somewhat doubtful value and I have therefore correlated goitre frequency with the cancer factor in this year for (i) the whole area, and (ii) all States east of Long. 104° W., and for the year 1920 for (iii) the 35 Registration States.

The actual mortality rates from cancer depend so greatly on the proportions of negroes to whites in the populations of the various States that they do not afford a true index of the real incidence of the disease, but the selection rates of cancer for any one organ (i.e. ratio of cancer of that organ to cancer of all organs) are less liable to distortion from this cause. I have only attempted to use actual death-rates for correlation in the case of the 1920 mortality data.

A second source of error is the variation in sex-ratio in the populations of States, the north-western States being characterized by a preponderance of males. Since cancer of the stomach forms a larger percentage of male cancers than of female cancers, the selection rate of cancer for the stomach in the whole population is positively correlated with the proportion of males. The mortality statistics do not separate the sexes and I have therefore attempted to eliminate this source of error by partial correlation in the 1900 data, whilst the 1920 mortality rates for cancer generally and for cancer of stomach and liver have been independently corrected for differing sex-distribution in the usual way before correlating.

A third source of error, which may also affect the selection rate as well as the actual death-rates, is differing age distribution, and an approximate standardization of the deaths has been carried out for 1920, again treating cancer of all organs and cancer of stomach and liver separately. In this way a selection rate for each State was obtained for 1920 which had been corrected for unclassified cancer deaths and also in a rough manner for differing age and sex distribution. For this year it was also possible to apply Professor Pearson's method of correlating actual numbers of cases for constant population in order to eliminate possible spurious correlation.

The principal cancer statistics used for obtaining these correlations are given in Table I, and the results are given below. For the sake of comparison correlations have also been worked out with the total tuberculosis rate in recruits, the diabetes rate in recruits, the diabetes death-rate in 1920, and the exophthalmic goitre rate in recruits.

(α) For the whole area*, correlating goitre rate in men with selection rate of cancer for the stomach obtained from 1900 statistics (correcting for cancer deaths where the organ was unspecified)

$$r = \cdot 509 \pm \cdot 073, \qquad \eta_{c.a} = \cdot 759.$$

This is clearly not a linear regression, but it becomes approximately so on omitting the 11 Western States with density of population less than 10 to the square mile, in which it is suspected that diagnosis was not so reliable (see β). The fact that the value of r then becomes larger suggests that there is a real relation which tends to be obscured by faulty statistics.

For comparison the following correlations over the same area may be noted.

Tuberculosis rate with goitre rate	$r = -\cdot 290 \pm \cdot 090$
" " with selection rate of cancer for stomach	$r = -\cdot 244 \pm \cdot 092$
Diabetes rate " " " " "	$r = \cdot 254 \pm \cdot 092$
" " " goitre rate	$r = \cdot 156 \pm \cdot 096$
" " " exophthalmic goitre rate	$r = \cdot 281 \pm \cdot 091$

(β) For the 37 States east of Long. 104° W., correlating goitre rate in men with selection rate of cancer for the stomach obtained from 1900 data as before

$$r = \cdot 644 \pm \cdot 065, \qquad \eta_{c.a} = \cdot 685.$$

* 47 areas, combining Arizona and New Mexico owing to their very small figures, and excluding Indian territory, Alaska and District of Columbia.

Since the regression now approximates to linearity I have endeavoured to eliminate the effect of differing sex distribution by partial correlation; and also to test whether the exophthalmic goitre correlation with the cancer factor arises from its own correlation with simple goitre.

Selection rate of cancer for stomach with percentage of males in population $r = \cdot534 \pm \cdot079$
Goitre rate with percentage of males ... ... $r = \cdot219 \pm \cdot105$
Selection rate with goitre, percentage of males being constant ... ... $r = \cdot639 \pm \cdot052$
Selection rate with exophthalmic goitre rate ... ... $\begin{cases} r = \cdot568 \pm \cdot075 \\ \eta_{c.g} = \cdot604 \end{cases}$
Goitre rate " " " " ... ... $\begin{cases} r = \cdot830 \pm \cdot034 \\ \eta_{g.c} = \cdot896 \end{cases}$
Exophthalmic goitre rate with percentage of males ... ... $r = \cdot287 \pm \cdot102$
Selection rate with exophthalmic goitre rate, percentage of males and simple goitre rate being constant ... ... $r = -\cdot031 \pm \cdot111$
Selection rate with goitre rate, percentage of males and exophthalmic goitre rate being constant ... ... $r = \cdot447 \pm \cdot089$

The last two figures suggest that the correlation between exophthalmic goitre and the cancer selection rate arises entirely from the relation between exophthalmic and simple goitre distributions. For comparison the following have been computed, percentage of males being constant.

Tuberculosis rate with selection rate of cancer for stomach $r = -\cdot268 \pm \cdot103$
Diabetes rate " " " " " $r = \cdot215 \pm \cdot106$
" " " goitre rate ... ... $r = \cdot168 \pm \cdot108$
" " " exophthalmic goitre rate ... ... $r = \cdot229 \pm \cdot105$

It may be noted that a positive relation between diabetes and total cancer death-rates for the American cities was demonstrated by Dr Maynard in this Journal (Vol. VII. p. 276)*. A significant relation between diabetes and exophthalmic goitre would be of interest in view of recent claims as to the beneficial effect of insulin treatment in Graves' Disease. If the diabetes death-rates for 1920 are correlated with those of exophthalmic goitre, however, $r = \cdot148 \pm \cdot108$, and with those of simple goitre $r = \cdot025 \pm \cdot111$, neither being significant.

(γ) For the 35 Registration States, correlating goitre rate in men with cancer mortality statistics for 1920, after correcting for age and sex distributions as well as for cancers of unspecified organ.

(i) Goitre cases, estimated, with deaths from cancer of stomach and liver, corrected ... ... $r = \cdot578 \pm \cdot076$
Ditto for population constant ... ... $r = \cdot438 \pm \cdot092$
(ii) Goitre rate with corrected mortality rate from cancer of stomach and liver for total cancer death-rate constant ... ... $r = \cdot516 \pm \cdot084$
(iii) Goitre rate with selection rate of cancer for stomach and liver ... ... $r = \cdot461 \pm \cdot090$

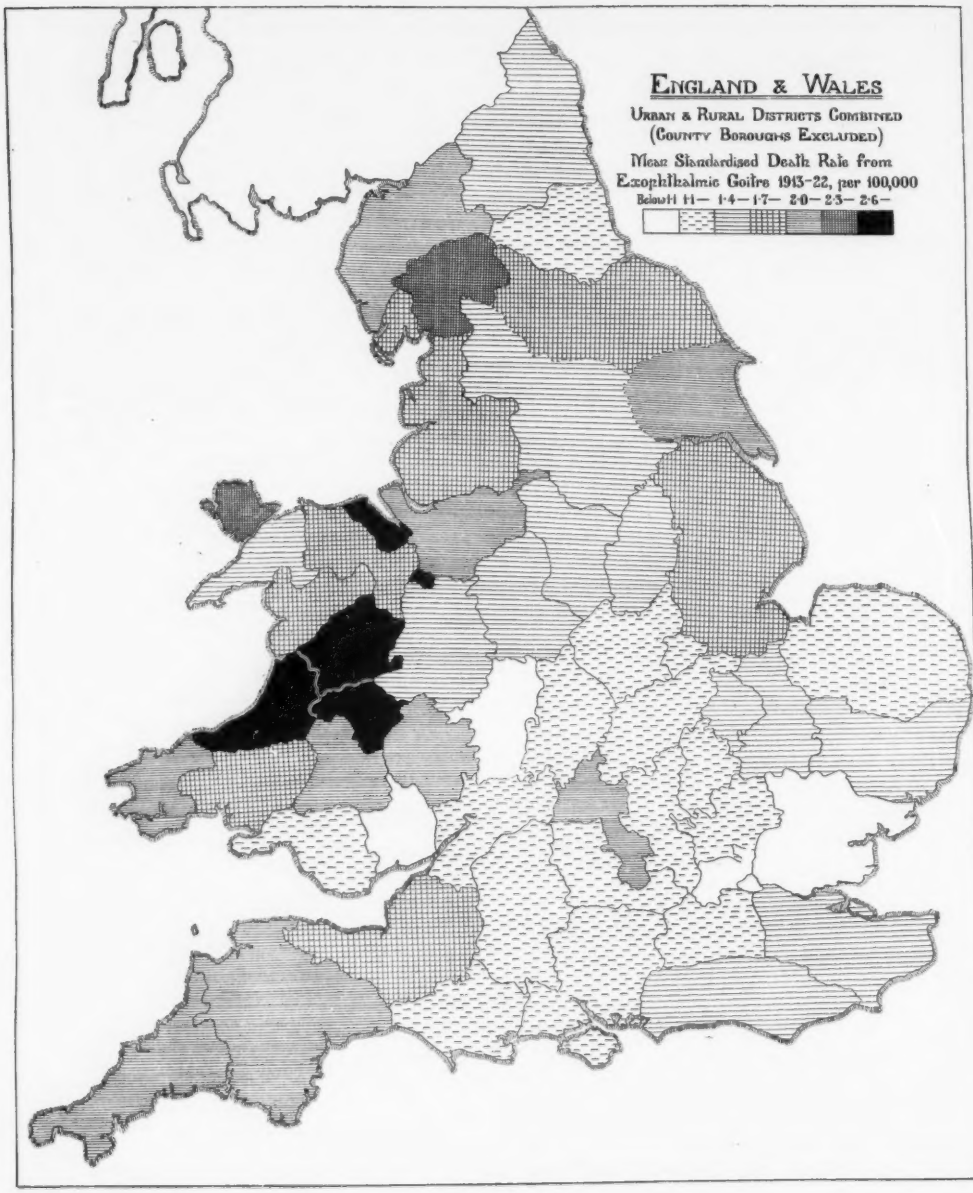
It will be noticed that the three methods of correlation give results of the same order.

These figures seem to indicate that there is evident in the American States as in Switzerland a significant relation between goitre frequency and the tendency of cancer to attack the stomach. In view of the more complex factors involved and the dubious value of the mortality statistics for some of the States this conclusion is put forward with due reserve, but the correspondence between the results obtained in such widely differing countries is remarkable.

II. *Goitre and Cancer in Sweden.*

Through the kind cooperation of Mr R. Moosberg of Stockholm I have been able to obtain statistics of investigations on goitre frequency in the two counties where goitre is most prevalent. In 1892 a survey of school-children aged 8–16, of confirmees aged 16–18 and of recruits for military service aged 21 and over was carried out in the Departments of Kopparberg and Gefleborg. In the former county as a whole 346 children out of 2221 examined were found to be

* For a criticism of Maynard's result: see *Biometrika*, Vol. XI. p. 191 *et seq.*


ENGLAND & WALES
URBAN & RURAL DISTRICTS COMBINED
(COUNTY BOROUGHS EXCLUDED)
Mean Standardised Death Rate from
Exophthalmic Goitre 1913-22, per 100,000
Below 1·1— 1·4— 1·7— 2·0— 2·3— 2·6—

MAP No. 2.

goitrous (15·6 %), 309 confirmees out of 1624 (19 %) and 109 soldiers out of 677 (16 %). A more recent survey in 1910 of the district of Husby in the same county showed 15·9 % of males and 33·6 % of females to have goitre. Other important endemic areas in the county were the districts of Bjursås, Sundborn, Stora Tuna and Svärisjo. In the Department of Gefleborg goitre was found to be endemic in a few areas, but for the county as a whole the percentage of recruits affected was only 1·3 % and of children 12·2 %. The most goitrous districts were Ockelbo with 44 % of children and 30 % of adolescents goitrous, and Bjenåker and Alfta with lower rates. A later investigation in the Högbo district showed 5 % of men and 30 % of women to be goitrous.

I have shown in Table II the total cancer deaths, deaths from all causes, deaths from unknown causes and corrected cancer mortality rates for the year 1911 for (i) the whole county of Kopparberg, (ii) goitrous areas of Ockelbo, Bjenåker and Alfta of the county of Gefleborg, and (iii) rural areas of Sweden as a whole.

TABLE II.

Area	Population 1911	Cancer deaths 1911	All deaths 1911	Deaths from unknown cause 1911	Corrected cancer death-rate
Department of Kopparberg ...	235,427	240	3067	240	—
Goitrous districts of Gefleborg	19,159	20	237	22	—
Total goitre areas	254,586	260	3304	262	110·9
Sweden (rural)	4,159,216	—	—	—	96·7

It will be seen that the goitrous areas had apparently a somewhat higher cancer mortality rate than rural Sweden as a whole. It is unfortunately not possible to obtain more extended data.

III. *Exophthalmic Goitre in England and Wales.*

In view of the interest which is at the present time being aroused on the question of goitre in England and Wales, I have thought it advisable to include with this paper a map (No. 2) showing the standardized mean mortality rates from exophthalmic goitre over the 10 years 1913–1922 in the various counties (excluding county boroughs) which has been compiled from the statistics of the Registrar-General. For explanation of the method used in standardizing reference should be made to Section VIII of the original paper in *Biometrika*, Vol. XVI. pp. 392—398, Dec. 1924.

Recurrence Formulae for the Moments of the Point Binomial.

By RAGNAR FRISCH, Kristiania.

Introduction. In *Biometrika*, May 1924, Professor Pearson has given a very important recurrence formula for the moments of the hypergeometric series. In the special case of the point binomial $(p+q)^r$ Pearson's formula may be written

$$\mu_s = rqp \sum_{i=0}^{s-2} \binom{s-1}{i} \mu_i - q \sum_{i=0}^{s-2} \binom{s-1}{i} \mu_{i+1} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (1)$$

or

$$\mu_s + q\binom{s-1}{1}\mu_{s-1} + \sum_{i=2}^{s} \left[q\binom{s-1}{i} - rqp\binom{s-1}{i-1} \right] \mu_{s-i} = 0 \quad \ldots\ldots\ldots\ldots\ldots\ldots (2),$$

where the moment of order s is defined to be

$$\mu_s = \mu_s(q) = \sum_{\nu=0}^{r} (\nu - rq)^s T_\nu,$$

where

$$T_\nu = \binom{r}{\nu} q^\nu p^{r-\nu}, \quad p+q=1 \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(3),$$

$$\mu_0 = 1,$$

$$\mu_1 = 0.$$

In the following lines I first deduce some general properties of linear equations between the moments of the point binomial. I then proceed to prove a system of recurrence formulae, a special case of which is the formula (1). Lastly I generalize the recurrence formulae of Pearson and Romanovsky* for the case of incomplete moments. The proofs of the last-named generalizations are based on principles entirely different from those originally used by Pearson and Romanovsky. In fact only elementary summation operations are involved. By the help of these operations the proofs may be given an extremely simple form. The formula of Romanovsky in fact arises with hardly any calculation at all.

§ 1. *Some General Properties of Linear Equations between the Moments of the Point Binomial.* A function $F(x)$ which identically in x satisfies one of the equations

$$F(1-x) = +F(x) \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(4a),$$

$$F(1-x) = -F(x) \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(4b),$$

may be called a *reduced* function of x, *positively* reduced if it satisfies $(4a)$, *negatively* reduced if it satisfies $(4b)$. A positively reduced function obviously is an even function in the variable $t=1-2x$, and a negatively reduced function is an odd function in the same variable.

Putting $p=1-q$ for q in (3) we obtain

$$\mu_s(1-q) = \sum_{\nu=0}^{r} (\nu - r + rq)^s \binom{r}{\nu} (1-q)^\nu q^{r-\nu} = (-1)^s \sum_{\nu=0}^{r} ((r-\nu) - rq)^s \binom{r}{r-\nu} q^{r-\nu} (1-q)^{r-(r-\nu)},$$

hence

$$\mu_s(1-q) = (-1)^s \mu_s(q) \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(5).$$

The even moments are positively reduced functions of q and the odd moments negatively reduced functions of q. That is: the even moments are even functions and the odd moments odd functions in the variable $\delta = p-q = 1-2q$.

Now let us consider a linear relation between the $(a+1)$ moments

$$\mu_s,\ \mu_{s-1},\ \ldots\ \mu_{s-a},$$

$$\sum_{i=0}^{a} A_i(q)\, \mu_{s-i} = 0 \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(6),$$

where the coefficients $A_i(q)$ are functions of q.

The function $\phi_i(q) = (-1)^i \cdot \dfrac{A_i(1-q)}{A_i(q)}$ may be called the *form-function* of the equation. If the form-function is independent of the subscript i, that is to say if a function $\phi(q)$ exists so that identically in q we have: $(-1)^i \cdot \dfrac{A_i(1-q)}{A_i(q)} = \phi(q)$ for all values $i=0, 1, \ldots a$, then I call (6) a *reduced* equation. If no such function exists, I call (6) a *not-reduced* equation.

The character of an equation as reduced or not-reduced evidently remains unaltered if the equation is multiplied by an arbitrary function of q.

* *Biometrika*, Vol. xv. p. 410, 1923. After finishing the present paper I received a letter from Professor Tchouproff who drew my attention to the fact that Romanovsky's formula was first found by Bohlmann. Cf. Bortkiewicz, *Jahresbericht der deutschen Mathematikervereinigung*, Bd. xxvii. (1918), S. 73. [Romanovsky's formula was communicated by him in 1915 to the Society of Naturalists of the University of Warsaw. Ed.]

In the special case $\phi(q) = \pm 1$ I call (6) a *completely reduced* equation, *positively* reduced if $\phi(q) = +1$, *negatively* reduced if $\phi(q) = -1$. A reduced equation which is not completely reduced is obviously transformed into a completely reduced equation when it is divided by any of its coefficients or any other function $F(q)$ which satisfies one of the equations

$$\phi(q) \cdot F(q) = \pm F(1-q),$$

where $\phi(q)$ is the form-function of the given equation.

Now suppose that we have given a not-reduced equation of the form (6). Since this equation holds good for every value of the variable between 0 and 1, it holds good for the value $(1-q)$. So we have

$$\sum_{i=0}^{a} A_i(1-q)\,\mu_{s-i}(1-q) = 0.$$

Substituting the expression for $\mu_{s-i}(1-q)$ from (5) we get

$$\sum_{i=0}^{a} (-1)^i A_i(1-q)\,\mu_{s-i} = 0 \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (7).$$

The last equation evidently is a not-reduced equation. Furthermore the two equations (6) and (7) are linearly independent since no function $\phi(q)$ exists, so that for all the values $i = 0, 1, \ldots a$ the coefficients $(-1)^i A_i(1-q)$ of the new equation (7) are $\phi(q)$ multiplied by the coefficients of the given equation (6).

Thus we have the following fundamental property of the moment equations: *To every not-reduced linear equation between the moments corresponds another not-reduced equation which is linearly independent of the first, and which may be written down immediately when the coefficients of the first equation are known.*

It is easily seen that this property does not belong to reduced equations. From the two not-reduced equations (6) and (7) we may by linear combination deduce an infinity of new equations. Of these however not more than two can be linearly independent. May it in this way be possible to obtain two linearly independent *reduced* equations?

Evidently it does not restrain generality if we suppose the equations to be obtained by adding to the equation (7) as it stands the equation (6) multiplied by some function $\theta(q)$ of q. This gives

$$\sum_{i=0}^{a} [\theta(q) A_i(q) + (-1)^i A_i(1-q)]\,\mu_{s-i} = 0 \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (8).$$

The form-function of this equation is

$$\Phi_i(q) = \frac{\theta(1-q)\,\phi_i(q) + 1}{\phi_i(q) + \theta(q)} = \theta(1-q)\,\frac{\phi_i(q) + \dfrac{1}{\theta(1-q)}}{\phi_i(q) + \theta(q)},$$

where $\phi_i(q)$ is the form-function of the given equation (6).

The necessary and sufficient condition that there exists a function $\Phi(q)$ independent of the subscript i, so that identically in q

$$\Phi_i(q) = \Phi(q)$$

for all the values $i = 0, 1, \ldots a$, is that

$$\theta(q)\,\theta(1-q) = 1 \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (8b),$$

for the values of q in the interval $0 < q < 1$ for which the form-function of (6), $\phi_i(q)$, is not independent of the subscript i. For the values of q for which (8*b*) is satisfied, we have

$$\Phi(q) = \theta(1-q) = \frac{1}{\theta(q)}.$$

Hence the necessary and sufficient condition that (8) is a reduced equation is that $\theta(q)$ is a function with the specified properties. Furthermore it is evident that if we choose two such functions $\theta_1(q)$ and $\theta_2(q)$ so that $\theta_1(q) \neq \theta_2(q)$, the corresponding two equations of the form (8) are linearly independent.

So we have the second property of the linear equations between the moments: *Any not-reduced equation of the form* (6) *may be replaced by two reduced linearly independent equations of the form*

$$\sum_{i=0}^{a} [\theta_1(q)\, A_i(q) + (-1)^i A_i(1-q)]\, \mu_{s-i} = 0 \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (9a),$$

$$\sum_{i=0}^{a} [\theta_2(q)\, A_i(q) + (-1)^i A_i(1-q)]\, \mu_{s-i} = 0 \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (9b),$$

where $\theta_1(q)$ *and* $\theta_2(q)$ *are two arbitrary not identical functions of* q *satisfying the conditions specified under* (8 *b*).

In treating linear equations between the moments one will nearly always find it easier to handle the reduced than the not-reduced forms.

Application. The equation (2) is a not-reduced equation with coefficients

$$A_0 = 1,$$

$$A_1(q) = q\binom{s-1}{1},$$

$$A_i(q) = q\binom{s-1}{i} - rqp\binom{s-1}{i-1}, \quad i = 2, 3, \ldots s.$$

Choosing $\theta_1(q) = +1$, $\theta_2(q) = -1$ we may replace the given equation by the following two reduced equations which are written down immediately by the help of (9*a*) and (9*b*),

$$2\mu_s = \sum_{i=1}^{\text{not v.}} \left[-(q + (-1)^i p)\binom{s-1}{i} + (1 + (-1)^i)\binom{s-1}{i-1} rqp \right] \mu_{s-i} \quad \ldots\ldots\ldots\ldots (10a),$$

$$\binom{s}{1}\mu_s = \sum_{i=1}^{\text{not v.}} \left[-(q + (-1)^i p)\binom{s}{i+1} + (1 + (-1)^i)\binom{s}{i} rqp \right] \mu_{s-i} \quad \ldots\ldots\ldots\ldots (10b),$$

where the summation is to be continued to the last not vanishing ("not v.") term. The equation (10*b*) is the equation obtained from (9*b*) when $(s-1)$ is replaced by s.

Writing out the formula (10*a*) we get

$$\mu_s = \binom{p-q}{2}\left[\binom{s-1}{1}\mu_{s-1} + \binom{s-1}{3}\mu_{s-3} + \ldots\right]$$
$$+ rqp\left[\binom{s-1}{1}\mu_{s-2} + \binom{s-1}{3}\mu_{s-4} + \ldots\right] - \frac{1}{2}\left[\binom{s-1}{2}\mu_{s-2} + \binom{s-1}{4}\mu_{s-4} + \ldots\right] \ldots\ldots (11),$$

where each of the sequences is to be continued to the last not vanishing term.

The formula (11) may be slightly more convenient for numerical calculations than the formula (1). Other formulae involving still less calculation may however be obtained as I shall show in the next paragraph.

§ 2. *A general System of Recurrence Formulae.* Let us denote by H_s the left-hand side of Pearson's equation (2),

$$H_s = \mu_s + q\binom{s-1}{1}\mu_{s-1} + \sum_{i=2}^{s}\left[q\binom{s-1}{i} - rqp\binom{s-1}{i-1}\right]\mu_{s-i}.$$

Calculating the first few differences of H_s,

$$\Delta H_s = H_{s+1} - H_s,$$
$$\Delta^2 H_s = \Delta H_{s+1} - \Delta H_s,$$
$$\ldots\ldots\ldots\ldots\ldots$$

the following formula for $\Delta^k H_s$ is suggested,

$$\Delta^k H_s = -\sum_{i=0}^{\text{not v.}} \left[(-1)^{i-1}\binom{k}{i} p - \binom{s-1}{i} q + \left(\binom{s-1}{i-1} + (-1)^i\binom{k}{i-1}\right) rqp \right] \mu_{s+k-i} \ldots (12).$$

The formula may easily be proved by complete induction. It obviously holds good for $k=0$. Supposing it to hold good for k we have

$$\begin{aligned}
&\Delta^{k+1}H_s=\Delta^k H_{s+1}-\Delta^k H_s\\
&=-\sum_{i=0}^{\text{not v.}}\left[(-1)^i\binom{k}{i}p-\binom{s}{i}q+\left(\binom{s}{i-1}+(-1)^i\binom{k}{i-1}\right)rqp\right]\mu_{s+1+k-i}\\
&\quad+\sum_{i=0}^{\text{not v.}}\left[(-1)^{i-2}\binom{k}{i-1}p-\binom{s-1}{i-1}q+\left(\binom{s-1}{i-2}+(-1)^{i-1}\binom{k}{i-2}\right)rqp\right]\mu_{s+k-(i-1)}\\
&=-\sum_{i=0}^{\text{not v.}}\left\{\left[(-1)^{i-1}\left(\binom{k}{i}+\binom{k}{i-1}\right)p-\left(\binom{s}{i}-\binom{s-1}{i-1}\right)q\right]\right.\\
&\qquad\left.+\left[\binom{s}{i-1}-\binom{s-1}{i-2}+(-1)^i\left(\binom{k}{i-1}+\binom{k}{i-2}\right)\right]rqp\right\}\mu_{s+(k+1)-i}\\
&=-\sum_{i=0}^{\text{not v.}}\left[(-1)^{i-1}\binom{k+1}{i}p-\binom{s-1}{i}q+\left(\binom{s-1}{i-1}+(-1)^i\binom{k+1}{i-1}\right)rqp\right]\mu_{s+(k+1)-i},
\end{aligned}$$

which is the expression for $\Delta^{k+1}H_s$ obtained from (12).

From Pearson's equation we know that $H_s=0$ identically in s, consequently $\Delta^k H_s=0$.

Hence replacing $(s+k)$ by s in (12) we have the following equation,

$$\sum_{i=0}^{\text{not v.}}\left[(-1)^{i-1}\binom{k}{i}p-\binom{s-k-1}{i}q+\left(\binom{s-k-1}{i-1}+(-1)^i\binom{k}{i-1}\right)rqp\right]\mu_{s-i}=0\ \ldots\ldots(13a),$$

where k denotes an arbitrary not negative integer. This is a not-reduced equation, which is easily seen by an inspection of the form-function of the equation*. Using the result of the preceding paragraph we may immediately write down the following not-reduced equation,

$$\sum_{i=0}^{\text{not v.}}\left[(-1)^{i-1}\binom{s-k-1}{i}p-\binom{k}{i}q+\left((-1)^i\binom{s-k-1}{i-1}+\binom{k}{i-1}\right)rqp\right]\mu_{s-i}=0\ldots(13b).$$

Taking first the sum and then the difference of the last two equations, which corresponds to choosing the functions $\theta_1(q)$ and $\theta_2(q)$ from the preceding paragraph equal to $+1$ and -1 respectively, we obtain the following two *reduced* equations,

$$\sum_{i=0}^{\text{not v.}}\left[\left(\binom{s-k-1}{i}+\binom{k}{i}\right)((-1)^{i-1}p-q)+(1+(-1)^i)\left(\binom{s-k-1}{i-1}+\binom{k}{i-1}\right)rqp\right]\mu_{s-i}=0$$
$$\ldots\ldots\ldots\ldots\ldots\ldots(14a),$$

$$\sum_{i=0}^{\text{not v.}}\left[\left(\binom{s-k}{i+1}-\binom{k}{i+1}\right)((-1)^{i-1}p-q)+(1+(-1)^i)\left(\binom{s-k}{i}-\binom{k}{i}\right)rqp\right]\mu_{s-i}=0\ \ldots(14b),$$

where k again denotes an arbitrary not negative integer, not necessarily the same in $(14a)$ as in $(14b)$. The equation $(14b)$ is the equation obtained by subtracting $(13a)$ from $(13b)$ and replacing $(s-1)$ by s.

To every not negative value of k correspond two recurrence formulae $(14a)$ and $(14b)$. Putting $k=0$ we obtain the formulae $(10a)$ and $(10b)$. In this special case the formulae contain all the moments from μ_0 to μ_s. By choosing k conveniently we may reduce the number of moments occurring in the formulae. In fact if we choose any particular value $k_0\left(<\frac{s}{2}\right)$ of k, the coefficients of all the moments from order zero to order $\begin{cases}k_0-1\text{ (if }(s-k_0)\text{ be even)}\\ k_0\text{ (if }(s-k_0)\text{ be odd)}\end{cases}$ vanish in the formula. If s be even and $k=\frac{s}{2}$, the coefficients in $(14a)$ vanish for the moments from order zero to order $\begin{cases}k-1\text{ (if }\frac{1}{2}s\text{ be even)}\\ k-2\text{ (if }\frac{1}{2}s\text{ be odd)}\end{cases}$. If s be odd we may for instance choose $k=\frac{1}{2}(s-1)$†. Formula

* The only case in which (13 *a*) might be a reduced equation is the case $k=\frac{1}{2}(s-1)$.

† $k=\frac{1}{2}(s-1)$ being the case in which (13 *a*) is a reduced equation, the special formula (15), p. 170, may of course also be derived directly from (13 *a*) or from (13 *b*).

(14a) then gives

$$\mu_s=(p-q)\left[\binom{k}{1}\mu_{s-1}+\binom{k}{3}\mu_{s-3}+\dots\right]+2rqp\left[\binom{k}{1}\mu_{s-2}+\binom{k}{3}\mu_{s-4}+\dots\right]$$
$$-\left[\binom{k}{2}\mu_{s-2}+\binom{k}{4}\mu_{s-4}+\dots\right] \quad \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots(15).$$

This formula involves considerably less calculation than the analogous formula (11). For instance to calculate μ_{13} by (11) we must use the values of all the moments from μ_0 to μ_{12} with numerical coefficients such as 495, 792, 924, etc. To calculate μ_{13} by (15) we only use the moments from μ_7 to μ_{12} with numerical coefficients 6, 15, and 20.

I think it should be possible by further transformation of Pearson's equation, using the results from § 1, to obtain an equation only containing four consecutive moments, but I have not yet been able to carry out this transformation.

§ 3. *Generalization of Pearson's and Romanovsky's Formulae to incomplete Moments.* The incomplete moment of order s may be defined as

$$\mu_s=\sum_{\nu=\rho}^{r}(\nu-rq)^s T_\nu,$$

where
$$T_\nu=\binom{r}{\nu}q^\nu p^{r-\nu} \quad \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots(16).$$

If it be desired to emphasize the lower limit of summation we may use the notation ${}_\rho\mu_s$.

Differentiating the equation (16) with respect to q we obtain

$$\mu_s'=\sum_{\nu=\rho}^{r}\left\{\frac{(\nu-rq)^{s+1}}{qp}T_\nu-rs(\nu-rq)^{s-1}T_\nu\right\},$$

hence
$$\mu_{s+1}=qp[rs\mu_{s-1}+\mu_s'] \quad \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots(17).$$

This I think is the simplest proof of which Romanovsky's formula is capable. It is a very interesting fact that the formula holds good without any alteration, even for the general case of incomplete moments.

For the incomplete moment of the first order $\mu_1=\sum_{\nu=\rho}^{r}(\nu-rq)T_\nu$ I have given the explicit expression in the *Skandinavisk Aktuarietidsskrift* *,

$$\mu_1=\rho p T_\rho=\rho p\binom{r}{\rho}q^\rho p^{r-\rho} \quad \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots(18),$$

Introducing this expression in the equation obtained from (17) by putting $s=0$, we get

$$\mu_0'=\frac{\rho}{q}T_\rho.$$

Integrating between 0 and q we obtain in the case $0<\rho$,

$$\mu_0=\rho\binom{r}{\rho}\int_0^q x^{\rho-1}(1-x)^{r-\rho}dx \quad \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots(19).$$

This is the expression for the zero moment of the point binomial first found by Professor Pearson †. r and ρ being positive integers, the integral (19) may of course be calculated directly, but the expression for μ_0 obtained would not be any simpler than the definition

$$\mu_0=\sum_{\nu=\rho}^{r}T_\nu \text{ ‡}.$$

The formula (19) may still be useful in numerical applications as shown by Dr Camp §. Now

* No. 3, 1924, p. 161. † *Biometrika*, May 1924, Vol. xvi. p. 202, and note by Dr Camp, p. 171.

‡ The most interesting results of the direct integration of (19) are obtained by equating coefficients of equal powers of q. In this way some not unimportant identities satisfied by the binomial coefficients may be obtained.

§ *Biometrika*, May 1924, Vol. xvi. p. 157.

assuming the numerical value of μ_0 to be calculated, I proceed to prove a recurrence formula by which the values of the higher incomplete moments may be calculated successively.

Putting
$$f_\nu=(\nu-rq)^{s-1},$$
$$g_\nu=(\nu-rq)\,T_\nu,$$

we have, by partial summation of (16),

$$\mu_s=\sum_{\nu=\rho}^{r} f_\nu g_\nu=f_r\sum_{\nu=\rho}^{r} g_\nu-\sum_{\nu=\rho}^{r-1}\Delta f_\nu\sum_{i=\rho}^{\nu} g_i.$$

Introducing the following expressions taken from (18),

$$\sum_{\nu=\rho}^{r} g_\nu=\rho p\,T_\rho,$$

$$\sum_{i=\rho}^{\nu} g_i=\rho p\,T_\rho-(\nu+1)\,p\,T_{\nu+1},$$

we get

$$\mu_s=\rho p T_\rho f_r-\rho p\,T_\rho\sum_{\nu=\rho}^{r-1}\Delta f_\nu+\sum_{\nu=\rho}^{r-1}p\,(\nu+1)\,T_{\nu+1}\Delta f_\nu=\rho p\,T_\rho f_\rho+q\sum_{\nu=\rho}^{r}(r-\nu)\,T_\nu\Delta f_\nu$$

$$=\rho p\,T_\rho f_\rho+q\sum_{\nu=\rho}^{r}(rq-\nu+rp)\,T_\nu\Delta f_\nu=\rho p\,T_\rho f_\rho+rqp\sum_{\nu=\rho}^{r}T_\nu\Delta f_\nu-q\sum_{\nu=\rho}^{r}(\nu-rq)\,T_\nu\Delta f_\nu.$$

Introducing $\quad \Delta f_\nu=\binom{s-1}{1}(\nu-rq)^{s-2}+\binom{s-1}{2}(\nu-rq)^{s-3}+\ldots+\binom{s-1}{s-1}$

we finally obtain

$$\mu_s=\rho p\,T_\rho(\rho-rq)^{s-1}+rqp\sum_{i=0}^{s-2}\binom{s-1}{i}\mu_i-q\sum_{i=0}^{s-2}\binom{s-1}{i}\mu_{i+1} \quad \ldots\ldots\ldots\ldots\ldots(20).$$

This is Pearson's formula generalized to the case of incomplete moments. The formula only differs from the complete-moment formula (1) by the additional term $\rho p\,T_\rho(\rho-rq)^{s-1}$. Putting $\rho=0$ in (20) we get of course (1). The incomplete moments do not satisfy any simple relation such as (5), so the equation (20) cannot be replaced by such simple reduced equations as (14*a*) and (14*b*).

The first few incomplete moments as calculated from (20) are

$$\mu_1=\rho p\,T_\rho,$$
$$\mu_2=\rho p\,T_\rho[\rho-(r+1)\,q]+rqp\,.\,\mu_0,$$
$$\mu_3=\rho p\,T_\rho[(\rho-(r+1)\,q)^2+qp\,(2r-1)]+rqp\,(p-q)\,.\,\mu_0,$$
$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$

Review: The Elements of Vital Statistics in their bearing on Social and Public Health Problems. By Sir ARTHUR NEWSHOLME, K.C.B., M.D., F.R.C.P. New Edition, George Allen and Unwin, Ltd.

This new edition of *The Elements of Vital Statistics* contains much fresh matter. Sir Arthur Newsholme discusses in an early chapter the different methods of estimating population. We think that he has some doubt as to the advantage of substituting "age in years and months" for "age last birthday" in the 1921 census. Personally we doubt whether it will conduce to accuracy. If anyone knows his age "age last birthday" seems to us the simplest form in which to state that age and one wonders how recorders deal with a portion of a month when they have to enter "age in years and months." In considering "age at marriage" Sir Arthur Newsholme shows that between 1896 and 1920 there has been very little postponement of marriage, ·88 of a year for bachelor bridegrooms and ·40 of a year for spinster brides, and one concludes that such a slight postponement can have little to do with the falling birth-rate.

The birth-rate is considered in Chapters VII, VIII, and IX. In connection with the registration of births we should like to point out that though the maiden surname of the mother is

required in Britain it is not always required when the birth of a British subject born abroad is registered and we think steps should be taken to remedy this defect in birth registration out of Britain where it occurs. We cannot see that there need be any connection between the sex ratios at birth and wholesale prices (p. 88) since both are varying with time. Later in the book there is a warning on this point but it is omitted in this instance.

As eugenists Sir Arthur Newsholme will not expect us to agree with much of what he writes on pp. 109, 110, and 111, though we fancy we see some "signs of grace" scattered here and there throughout the book and even on the pages referred to above! In the discussion on Infant Mortality it is assumed that a lower infant death-rate means improved health in the survivors; we still dare to think that this is the expression of a pious hope which is at present unproven. The treatment of life-tables is very superficial*.

We wish that heed would be taken of the chapter on the Registration of Sickness; we admit the difficulties but they should not be insuperable and the same seems to us to apply to the problem of death certificates though we think there will always be a difficulty over international statistics, even if we in England adopted the Swiss system. We notice on p. 290 that "Nothing is more certain than the statement that the removal of poverty would effect an enormous reduction of disease"; and in the next paragraph but one a quotation from the *Report of the Royal Commission on the Poor Laws* (1909) which states that "Sickness is therefore admittedly one of the chief causes of pauperism"; these two quotations seem to us in opposition and both require investigation before anyone can say which is the truer statement. Later on when it comes to the question of infant mortality Sir Arthur Newsholme points out the difficulty of disentangling cause and effect when poverty and sickness occur side by side; but though he has pointed this out he does not hesitate to state that "Environment has a preponderant influence in determining the magnitude of infantile and child mortality in a community"; by environment is meant, we presume, poverty, housing conditions, &c. Material collected by medical officers of health in individual districts seems to show that the most important factors are the health of the parents and the care that they bestow on their children, and though we consider it not yet proven, we are inclined to think that the fall in the infant death-rate of late years is due to the Midwives' Act and possibly, though hitherto there is no statistical proof, to some improvement in the habits of the mothers.

Again we cannot see that there need be any connection between prices and phthisis or that it is necessary to try to explain differences between England and Ireland in this connection. Sir Arthur Newsholme does not assume that there is a close connection between prices and tuberculosis but we think that to give coefficients of correlation when there is a time factor introduced is very dangerous unless they are by one method or another corrected for time. It would be interesting to know whether the number of people attacked by phthisis had diminished or not but such information would be, we think, impossible to obtain, since tuberculosis is now diagnosed at an earlier stage than formerly. The old question of the influence of heredity in the development of phthisis comes up again; some people among whom we count ourselves still believe in an hereditary predisposition but if one assumes that the association between tuberculosis in parent and child is due to large doses of infection at an early age rather than to an hereditary predisposition the argument against the phthisical having children seems to us equally strong; the children will be more likely to have tuberculosis than the children of parents without the disease.

In a review in *Biometrika* the statements in connection with the *Variate Difference Method* on p. 524 cannot be ignored. Sir Arthur Newsholme has quoted from a paper by Prof. Person of Harvard in which Prof. Person says that "the defective character of the variate difference method, as originally proposed, is admitted." We think that Sir Arthur Newsholme has not read the paper

* The result on p. 238 for p_x has been obtained from *English Life Tables*, 1914, p. 60, without Sir Arthur Newsholme troubling to test his own arithmetic. Had he done so, he would, perhaps, better have understood the meaning of the ordinary life-table symbols.

in *Biometrika* to which Prof. Person refers. There was an error in the application of the method to the problem of infant death-rate which the authors of that paper admitted but that is a very different matter from an error in the method itself. The product-moment method of obtaining a correlation coefficient is perfectly sound but it may be applied without taking into account the time factor or that regression is non-linear, and erroneous conclusions may be drawn from the coefficient of correlation found; such applications do not invalidate the method, and it has yet to be proved that the variate difference method is fundamentally at fault. We cannot help wishing that Sir Arthur Newsholme had ended his book at p. 500. There are many useful notes and "cautions" in the last 100 pages but also very much that seems to us open to criticism. We think the root of the difference between Sir Arthur Newsholme and the biometricians is to be found on p. 525 where it is suggested that "when the conclusions drawn from coefficients of correlation run counter to inferences on the same subject based on biological (including medical) considerations" judgment should be suspended. This seems to the statistician a most dangerous course to pursue. That a careful investigation of the results obtained by statistical methods is necessary we should all agree, but when this has been done we feel that the results must be accepted and that to accept only those results which agree with existing prejudgments is a retrograde step which must lead to disaster. We think it is this underlying idea which leads Sir Arthur Newsholme to reject evidence provided in the case of two collections of material and accept results deduced from another far more scanty and incomplete. On p. 558 the following sentence occurs: "Another example, on a somewhat scanty scale, may be quoted from McDougall's book. Miss Aylitt tested 342 children from the primary grades of schools in one district," but there is no reference to scantiness of material when the intelligence of 372 girls in one school and of 247 boys in another are tested and far-reaching conclusions are drawn by Dr Isserlis from this material. Sir Arthur Newsholme assumes that the teachers in the two schools referred to were really testing intelligence and not knowledge and one infers that he thinks that in all previous investigations it was the other way round! The correlations found with age make one doubtful and the comparatively low association between the teacher's estimate of intelligence and intelligence judged by a test makes one still more doubtful, but the numbers, reduced in these tables to 161 girls and 106 boys, are quite inadequate. We fear that it is because the conclusions reached do run counter to previous inferences on the subject that they are accepted uncritically.

This book gives an excellent summary of what material is available in the census returns and shows how the statistics in this country could be improved in many ways, as for instance by the addition of the age of the two parents when the birth of an infant is registered, but we feel obliged to object to many of the conclusions, considering that some are contrary to existing evidence and that others are quite unproven. The book is bound to be replaced at no distant date, by one on sounder lines. E. M. E.

Lebenserwartung und mittleres Alter der Lebenden.

Von E. I. GUMBEL, Heidelberg.

Im folgenden sollen die Grössenbeziehungen zwischen der zu einem bestimmten Alter gehörigen Lebenserwartung und dem mittleren Alter der nach der Sterbetafel über diesem Alter Lebenden untersucht werden. Für das Alter Null ist bereits in einer früheren Arbeit* gezeigt, dass für alle heutigen Sterbetafeln die Lebenserwartung eines Neugeborenen grösser ist, als das mittlere Alter aller Lebenden. Ein Ansatz zur Erweiterung dieses Satzes findet sich in einer Arbeit von Pearson †, der gezeigt hat, dass bei der Gompertz-Makehamschen Absterbeordnung für jedes Alter die Summe Alter plus Lebenserwartung grösser ist als das hierzu gehörige mittlere Alter der Lebenden. Es soll nun untersucht werden, unter welchen allgemeinen Bedingungen dieser Satz gilt. Ferner soll auch eine untere Grenze für das mittlere Alter der Lebenden aufgestellt werden.

* "Uber eine Eigenschaft der Sterbetafel," *Zeitschrift für die gesamte Versicherungswissenschaft*, Bd. XXIV. H. 4, 1924.

† *Biometrika*, Vol. XVI. p. 297, 1924.

Man bezeichnet mit $l(x)$ die Wahrscheinlichkeit, die einem Neugeborenen zukommt, den $x$ten Geburtstag zu erreichen, sodass $l(0)=1$. Man definiere das höchste erreichbare Alter $\omega-1$ durch die zweite Grenzbedingung $l(\omega)=0$. Ferner definiere man die Sterbensintensität $\mu(x)$ wie üblich, als

$$\mu(x)=-\frac{d\,l(x)}{l(x)\,dx}.$$

Das mittlere Alter für die nach der Sterbetafel über dem Alter x Lebenden lautet

$$\overline{x(x)}=\frac{\int_x^\omega y\,l(y)\,dy}{\int_x^\omega l(y)\,dy}.$$

Die Integrationsvariable wird hier stets mit y bezeichnet, um Verwechselungen mit der unteren Grenze x zu vermeiden. Die Lebenserwartung wird wie üblich definiert als

$$E(x)=\frac{1}{l(x)}\int_x^\omega l(y)\,dy.$$

Man bezeichne die Summe Alter + Lebenserwartung als Alterserwartung.

Somit gilt für die Gompertz-Makehamsche Absterbeordnung nach Pearson

$$\overline{x(x)}<x+E(x).$$

Dieser Satz lässt sich nach einem von Herrn v. Bortkiewicz* stammenden Verfahren für alle Absterbeordnungen beweisen, für welche die Sterbensintensität stets mit dem Alter wächst. Zum Beweise sei zunächst eine einfache Beziehung abgeleitet, die im folgenden von nöten ist. Man betrachte, nach der Definition der Sterbensintensität,

$$\int_x^\omega y^m\mu(y)\,l(y)\,dy=-\int_x^\omega y^m\,d\,l(y)$$

$$=x^m l(x)+m\int_x^\omega y^{m-1}l(y)\,dy.$$

Im Spezialfall

$m=0$: wird $\int_x^\omega \mu(y)\,l(y)\,dy=l(x)$,

$m=1$: wird $\int_x^\omega y\,\mu(y)\,l(y)\,dy=x\,l(x)+\int_x^\omega l(y)\,dy=l(x)\,(x+E(x))$,

$m=2$: wird $\int_x^\omega y^2\mu(y)\,l(y)\,dy=x^2 l(x)+\frac{2\int_x^\omega y\,l(y)\,dy}{\int_x^\omega l(y)\,dy}\int_x^\omega l(y)\,dy=l(x)\,(x^2+2\overline{x(x)}\,E(x))$.

Unsere Aufgabe ist nun, Alterserwartung und mittleres Alter zu vergleichen. Man bezeichne

$$\sigma=\operatorname{sign}[x+E(x)-\overline{x(x)}],$$

$$=\operatorname{sign}[l(x)\,(x+E(x))-l(x)\,\overline{x(x)}],\ \text{da } l(x)>0,$$

$$=\operatorname{sign}\left[\int_x^\omega (y-\overline{x(x)})\,\mu(y)\,l(y)\,dy\right]$$

nach den eben für $m=0$ und $m=1$ abgeleiteten Formeln.

Ferner ist: $\sigma=\operatorname{sign}\left[\mu(\phi)\int_x^{\overline{x(x)}}(y-\overline{x(x)})\,l(y)\,dy+\int_{\overline{x(x)}}^\omega (y-\overline{x(x)})\,\mu(y)\,l(y)\,dy\right]$

* L. v. Bortkiewicz, *Die mittlere Lebensdauer*, Jena, G. Fischer, 1893, p. 77.

nach dem Mittelwertsatz der Integralrechnung, wobei $x < \phi < \overline{x(x)}$. Aber nach der Definition des mittleren Alters der Lebenden ist

$$\int_x^{\overline{x(x)}} (y - \overline{x(x)})\, l(y)\, dy = -\int_{\overline{x(x)}}^{\omega} (y - \overline{x(x)})\, l(y)\, dy,$$

sodass

$$\sigma = \operatorname{sign}\left[\int_{\overline{x(x)}}^{\omega} (y - \overline{x(x)})\,(\mu(y) - \mu(\phi))\, l(y)\, dy\right]$$
$$= \operatorname{sign} \mu'(x)$$

da wegen der Integrationsgrenzen $\phi < \overline{x(x)} < y$ und da $l(x)$ stets positiv.

Also gilt: *Die Alterserwartung ist grösser, gleich oder kleiner als das mittlere Alter der Lebenden, wenn die Sterbensintensität für alle Alter wächst, gleich bleibt oder fällt.* Dieser Satz lässt sich aber nicht umkehren. Es kann sehr wohl die Alterserwartung grösser sein als das mittlere Alter der Lebenden, ohne dass die Sterbensintensität stets wächst.

Der Pearsonsche Satz ergibt sich als Spezialfall. Denn die Gompertz-Makehamsche Formel geht aus von

$$\mu(x) = A + Bc^x,$$

woraus

$$\mu'(x) = Bc^x \log c < 0,$$

weil

$$B > 0,$$

und

$$c > 1.$$

Somit haben wir eine unter einer Bedingung gültige obere Grenze für das mittlere Alter der Lebenden aufgestellt. Nach einem allgemein für solche Abschätzungen verwandten Verfahren ist es leicht, auch eine untere Grenze hierfür aufzustellen. Man bilde die Wahrscheinlichkeit dafür, dass ein Neugeborener den $x$ten Geburtstag nicht erreicht,

$$g(x) = 1 - l(x),$$

dann gilt für die Momente dieser Funktion

$$\int_x^{\omega} y^m\, dg(y) = -\int_x^{\omega} y^m\, dl(y).$$

Man bilde nun den stets positiven Ausdruck $A = \frac{1}{l(x)} \int_x^{\omega} (a_1 + a_2 y)^2\, dg(y)$, wobei a_1 und a_2 vorläufig willkürliche reelle Konstanten sind. Führt man die Integration aus, so wird, wenn man die oben berechneten Momente verwendet,

$$A = a_1^2 + 2a_1 a_2 (x + E(x)) + a_2^2 (x^2 + 2\overline{x(x)}\, E(x)) > 0.$$

Nimmt man die quadratische Ergänzung vor, so wird

$$(a_1 + a_2 (x + E(x))^2 > a_2^2 (x + E(x))^2 - a_2^2 (x^2 + 2\overline{x(x)}\, E(x)).$$

Denkt man sich jedes einzelne Alter speziell betrachtet und setzt man hierfür jeweils

$$a_1 = -a_2 (x + E(x)),$$

so wird

$$x^2 + 2\overline{x(x)}\, E(x) > x^2 + 2x\, E(x) + E(x)^2, \text{ da } a_2^2 > 0,$$

oder

$$\overline{x(x)} > x + \frac{E(x)}{2}.$$

Demnach gilt der Satz: *Das mittlere Alter der nach der Sterbetafel über dem Alter x Lebenden ist grösser als das Alter x+der halben Lebenserwartung.* Für das Alter null gilt speziell: das mittlere Alter aller nach der Sterbetafel Lebenden ist grösser als die halbe Lebenserwartung eines Neugeborenen. Somit ist es gelungen, das mittlere Alter der Lebenden mit Hilfe der Lebenserwartung in feste Grenzen einzuschliessen.

FURTHER CONTRIBUTIONS TO THE THEORY OF SMALL SAMPLES*.

By KARL PEARSON, F.R.S.

I. *On the Correlation of the two Standard Deviations and of either of these with the Coefficient of Correlation in small Samples taken from a bi-variate Population following the Normal Law of Distribution.*

The distribution of the two standard deviations and the coefficient of correlation, namely σ_1, σ_2 and r, in samples of n taken from a normal population of constants Σ_1, Σ_2 and ρ is known to be†:

$$z = z_0 e^{-\frac{1}{2(1-\rho^2)}\left(\frac{n\sigma_1^2}{\Sigma_1^2} - \frac{2nr\rho\sigma_1\sigma_2}{\Sigma_1\Sigma_2} + \frac{n\sigma_2^2}{\Sigma_2^2}\right)} \left(\frac{\sigma_1\sigma_2}{\Sigma_1\Sigma_2}\right)^{n-2} (1-r^2)^{\frac{n-4}{2}} \quad \text{.....(i).}$$

Let us write

$$s_1 = \Sigma_1\sqrt{1-\rho^2}/\sqrt{n}, \quad s_2 = \Sigma_2\sqrt{1-\rho^2}/\sqrt{n}, \quad n\rho\sigma_1\sigma_2/\Sigma_1\Sigma_2(1-\rho^2) = \rho\sigma_1\sigma_2/s_1s_2;$$

then

$$z = z_0' e^{-\frac{1}{2}\frac{\sigma_1^2}{s_1^2}} e^{-\frac{1}{2}\frac{\sigma_2^2}{s_2^2}} \left(\frac{\sigma_1\sigma_2}{s_1s_2}\right)^{n-2} e^{\gamma r} (1-r^2)^{\frac{n-4}{2}} \quad \text{.....(ii),}$$

where z_0' differs from z_0 by terms not containing σ_1, σ_2 or r, and $\gamma = \rho\sigma_1\sigma_2/s_1s_2$.

Accordingly the frequency surface for σ_1, σ_2 will be obtained by integrating r throughout its entire range, i.e. from -1 to $+1$. We require accordingly the integral

$$I = \int_{-1}^{+1} e^{\gamma r} (1-r^2)^{\frac{n-4}{2}} dr \quad \text{.....(iii)}$$

Take $r = \sin\theta$, and we have:

$$I = \int_{-\frac{\pi}{2}}^{+\frac{\pi}{2}} e^{\gamma\sin\theta} \cos^{n-3}\theta d\theta.$$

Expand $e^{\gamma\sin\theta}$ and note that all terms

$$\int_{-\frac{\pi}{2}}^{+\frac{\pi}{2}} \sin^{2p+1}\theta\cos^{n-3}\theta d\theta = 0.$$

Accordingly

$$I = 2\int_0^{\frac{\pi}{2}} \left(1 + \frac{\gamma^2}{2!}\sin^2\theta + \frac{\gamma^4}{4!}\sin^4\theta + \ldots\right)\cos^{n-3}\theta d\theta.$$

* Printed from Lecture Notes.

† R. A. Fisher, *Biometrika*, Vol. x. p. 510, with a different notation.

Put $x = \sin^2\theta$ and we have

$$I = \int_0^1 \left(x^{-\frac{1}{2}}(1-x)^{\frac{n-4}{2}} + \frac{\gamma^2}{2!}x^{\frac{1}{2}}(1-x)^{\frac{n-4}{2}} + \frac{\gamma^4}{4!}x^{\frac{3}{2}}(1-x)^{\frac{n-4}{2}} + \ldots + \frac{\gamma^{2p}}{2p!}x^{\frac{2p-1}{2}}(1-x)^{\frac{n-4}{2}} + \ldots \right) dx$$

$$= B\left(\frac{1}{2}, \frac{n-2}{2}\right) + \frac{\gamma^2}{2!}B\left(\frac{3}{2}, \frac{n-2}{2}\right) + \frac{\gamma^4}{4!}B\left(\frac{5}{2}, \frac{n-2}{2}\right) + \ldots + \frac{\gamma^{2p}}{2p!}B\left(\frac{2p+1}{2}, \frac{n-2}{2}\right) + \ldots \quad \ldots\ldots\text{(iv)},$$

where $B(p, q)$ is the usual beta-function,

$$= B\left(\frac{1}{2}, \frac{n-2}{2}\right)\left(1 + \frac{\gamma^2}{2!}\frac{1}{n-1} + \frac{\gamma^4}{4!}\frac{3.1}{(n-1)(n+1)} + \ldots + \frac{\gamma^{2p}}{2p!}\frac{(2p-1)(2p-3)\ldots 1}{(n-1)(n+1)\ldots(n+2p-3)} + \ldots\right)$$

$$= B\left(\frac{1}{2}, \frac{n-2}{2}\right)\left(1 + \frac{\gamma^2}{1!}\cdot\frac{1}{2n-2} + \frac{\gamma^4}{2!}\cdot\frac{1}{(2n-2)(2n+2)} + \frac{\gamma^6}{2!}\cdot\frac{1}{(2n-2)(2n+2)(2n+6)} + \ldots + \frac{\gamma^{2p}}{p!}\cdot\frac{1}{(2n-2)(2n+2)\ldots(2n+4p-6)} + \ldots\right).$$

We can now return to (ii) and write down our equation to the frequency surface for σ_1, σ_2. It is

$$y = y_0 e^{-\frac{1}{2}\frac{\sigma_1^2}{s_1^2}} e^{-\frac{1}{2}\frac{\sigma_2^2}{s_2^2}} \left(\frac{\sigma_1\sigma_2}{s_1 s_2}\right)^{n-2}\left(1 + \frac{\rho^2}{1!}\frac{\sigma_1^2\sigma_2^2}{s_1^2 s_2^2}\frac{1}{2n-2} + \frac{\rho^4}{2!}\frac{\sigma_1^4\sigma_2^4}{s_1^4 s_2^4}\frac{1}{(2n-2)(2n+2)} + \ldots + \frac{\rho^{2p}}{p!}\frac{\sigma_1^{2p}\sigma_2^{2p}/s_1^{2p}s_2^{2p}}{(2n-2)(2n+2)\ldots(2n+4p-6)} + \ldots\right) \quad \ldots\ldots\ldots\ldots\text{(v)}.$$

It is clear that in order to find any product moment function of this surface *about the* zero of σ_1, σ_2 we need only to find the value of

$$I' = \int_0^\infty e^{-\frac{1}{2}\frac{\sigma_1^2}{s_1^2}}\left(\frac{\sigma_1}{s_1}\right)^{2p}\left(\frac{\sigma_1}{s_1}\right)^m d\sigma_1$$

$$= s_1 \times \int_0^\infty e^{-\frac{1}{2}\frac{\sigma_1^2}{s_1^2}}\left(\frac{\sigma_1}{s_1}\right)^{2p+m} d\frac{\sigma_1}{s_1}.$$

Write $$z = \frac{1}{2}\frac{\sigma_1^2}{s_1^2} \quad \text{or} \quad \frac{\sigma_1}{s_1} = \sqrt{2z}.$$

Hence I' is given by

$$I' = 2^{2p+m-1} s_1^{m+1} \int_0^\infty e^{-z} z^{\frac{2p+m-1}{2}} dz$$

$$= 2^{\frac{2p+m-1}{2}} s_1 \times \Gamma\left(\frac{2p+m+1}{2}\right) \quad \ldots\ldots\ldots\ldots\text{(vi)}.$$

If N be the total number of samples, we have:

$$N = \int_0^\infty \int_0^\infty y\, d\sigma_1 d\sigma_2$$

$$= y_0 2^{n-3}(s_1 s_2) \times \left\{\Gamma^2\left(\frac{n-1}{2}\right) + \rho^2 2^2 \Gamma^2\left(\frac{n+1}{2}\right)\frac{1}{1!(2n-2)}\right.$$

$$+ \rho^4 2^4 \Gamma^2\left(\frac{n+3}{2}\right)\frac{1}{2!(2n-2)(2n+2)} + \ldots$$

$$\left. + \rho^{2p} 2^{2p} \Gamma^2\left(\frac{2p+m+1}{2}\right)\frac{1}{p!(2n-2)(2n+2)\ldots(2n+4p-6)} + \ldots\right\}$$

$$= y_0 2^{n-3}(s_1 s_2) \times \Gamma^2\left(\frac{n-1}{2}\right)\left(1 + \frac{n-1}{2}\frac{\rho^2}{1!} + \frac{n-1}{2}\frac{n+1}{2}\frac{\rho^4}{2!} + \ldots\right.$$

$$\left. + \frac{n-1}{2}\frac{n+1}{2}\ldots\frac{n-1+2p}{2}\frac{\rho^{2p}}{p!} + \ldots\right).$$

Or
$$N = y_0 \frac{2^{n-3}(s_1 s_2) \times \Gamma^2\left(\frac{n-1}{2}\right)}{(1-\rho^2)^{\frac{n-1}{2}}} \quad \ldots\ldots\ldots\ldots\ldots\ldots\text{(vii).}$$

This determines y_0, i.e.

$$y_0 = \frac{N(1-\rho^2)^{\frac{n-1}{2}}}{2^{n-3}(s_1 s_2) \times \Gamma^2\left(\frac{n-1}{2}\right)} \quad \ldots\ldots\ldots\ldots\ldots\ldots\text{(viii).}$$

We shall now find

$$Np'_{\sigma_1\sigma_2} = \int_0^\infty \int_0^\infty y\,\sigma_1\sigma_2\, d\sigma_1 d\sigma_2$$

$$= s_1^2 s_2^2 \int_0^\infty \int_0^\infty y \frac{\sigma_1}{s_1}\frac{\sigma_2}{s_2} d\frac{\sigma_1}{s_1} d\frac{\sigma_2}{s_2}$$

$$= y_0 s_1^2 s_2^2\, 2^{n-2}\left\{\Gamma^2\left(\frac{n}{2}\right) + \frac{\rho^2}{1!}2^2\Gamma^2\left(\frac{n+2}{2}\right)\frac{1}{2n-2}\right.$$

$$+ \frac{\rho^4}{2!}2^4\Gamma^2\left(\frac{n+4}{2}\right)\frac{1}{(2n-2)(2n+2)} + \ldots$$

$$\left. + \frac{\rho^{2p}}{p!}2^{2p}\Gamma^2\left(\frac{n+2p}{2}\right)\frac{1}{(2n-2)(2n+2)\ldots(2n+4p-6)} + \ldots\right\}$$

$$= y_0 s_1^2 s_2^2\, 2^{n-2}\Gamma^2\left(\frac{n}{2}\right)\left(1 + \frac{\frac{1}{2}n\,\frac{1}{2}n}{\frac{1}{2}(n-1).1!}\rho^2\right.$$

$$+ \frac{\frac{1}{2}n(\frac{1}{2}n+1)\,\frac{1}{2}n(\frac{1}{2}n+1)}{\frac{1}{2}(n-1)\,\frac{1}{2}(n+1).2!}\rho^4 + \ldots$$

$$\left. + \frac{\frac{1}{2}n(\frac{1}{2}n+1)\ldots\frac{1}{2}(n+p)\,\frac{1}{2}n(\frac{1}{2}n+1)\ldots\frac{1}{2}(n+p)}{\frac{1}{2}(n-1)\,\frac{1}{2}(n+1)\ldots\frac{1}{2}(n+2p-3).p!}\rho^{2p} + \ldots\right)$$

$$= y_0 s_1^2 s_2^2\, 2^{n-2}\Gamma^2\left(\frac{n}{2}\right)F\left(\tfrac{1}{2}n,\ \tfrac{1}{2}n,\ \tfrac{1}{2}(n-1),\ \rho^2\right),$$

where $F(\alpha, \beta, \gamma, \rho^2)$ is the usual notation for the hypergeometrical series. But by Euler's Theorem:

$$F(\tfrac{1}{2}n, \tfrac{1}{2}n, \tfrac{1}{2}(n-1), \rho^2) = (1-\rho^2)^{-\frac{1}{2}(n+1)} F(\tfrac{1}{2}, \tfrac{1}{2}, \tfrac{1}{2}(n-1), \rho^2).$$

Accordingly we have:

$$Np'_{\sigma_1\sigma_2} = y_0 s_1^2 s_2^2 2^{n-2} \Gamma^2\left(\frac{n}{2}\right)(1-\rho^2)^{-\frac{1}{2}(n+1)} \times H(\rho^2),$$

where

$$H(\rho^2) = \left(1 + \frac{\rho^2}{1!}\frac{1}{2n-2} + \frac{\rho^4}{2!}\frac{1}{(2n-2)(2n+2)} + \cdots + \frac{\rho^{2p}}{p!}\frac{1}{(2n-2)(2n+2)\ldots(2n+4p-6)} + \ldots\right) \ldots\text{(ix)}.$$

Accordingly

$$p'_{\sigma_1\sigma_2} = \frac{2s_1 s_2}{1-\rho^2}\frac{\Gamma^2\left(\frac{n}{2}\right)}{\Gamma^2\left(\frac{n-1}{2}\right)} H(\rho^2) \ldots\ldots\ldots\ldots\ldots\ldots\text{(x)}.$$

Returning to Equation (v) I now proceed to integrate it for one value σ_2 of the two variates only. We have

$$\int_0^\infty y\, d\sigma_2 = y_0 e^{-\frac{1}{2}\frac{\sigma_1^2}{s_1^2}}\left(\frac{\sigma_1}{s_1}\right)^{n-2} \times s_2$$

$$\times \int_0^\infty e^{-\frac{1}{2}\frac{\sigma_2^2}{s_2^2}}\left(\frac{\sigma_2}{s_2}\right)^{n-2} \sum_0^\infty \frac{\rho^{2p}}{p!}\frac{\sigma_1^{2p}\sigma_2^{2p}}{s_1^{2p}s_2^{2p}}\frac{1}{(2n-2)(2n+2)\ldots(2n+4p-6)}\, d\left(\frac{\sigma_2}{s_2}\right)$$

$$= y_0 s_2 e^{-\frac{1}{2}\frac{\sigma_1^2}{s_1^2}}\left(\frac{\sigma_1}{s_1}\right)^{n-2} \times \sum_0^\infty \frac{\rho^{2p}}{p!}\left(\frac{\sigma_1}{s_1}\right)^{2p}$$

$$\times 2^{\frac{2p+n-3}{2}}\Gamma\left(\frac{n-1}{2}+p\right)\frac{1}{(2n-2)(2n+2)\ldots(2n+4p-6)}$$

$$= y_0 s_2 2^{\frac{n-3}{2}}\Gamma\left(\frac{n-1}{2}\right) e^{-\frac{1}{2}\left(\frac{\sigma_1}{s_1}\right)^2}\left(\frac{\sigma_1}{s_1}\right)^{n-2}\left\{1 + \frac{\rho^2}{1!}\left(\frac{\sigma_1}{s_1}\right)^2\frac{n-1}{2(n-1)}\right.$$

$$\left. + \frac{\rho^4}{2!}\left(\frac{\sigma_1}{s_1}\right)^2\frac{(n-1)(n-2)}{2^2(n-1)(n-2)} + \ldots\right\}$$

$$= y_0 s_2 2^{\frac{n-3}{2}}\Gamma\left(\frac{n-1}{2}\right)\left(\frac{\sigma_1}{s_1}\right)^{n-2} e^{-\frac{1}{2}\frac{\sigma_1^2}{s_1^2}+\frac{1}{2}\rho^2\frac{\sigma_1^2}{s_1^2}}.$$

But $s_1 = \frac{\Sigma_1}{n}\sqrt{1-\rho^2}$; hence substituting we have for the frequency distribution:

$$y = \frac{y_0 s_2 2^{\frac{n-3}{2}} n^{\frac{n-2}{2}}}{(1-\rho^2)^{\frac{n-2}{2}}} e^{-\frac{1}{2}\frac{n\sigma_1^2}{\Sigma_1^2}}\left(\frac{\sigma_1}{\Sigma_1}\right)^{n-2}\Gamma\left(\frac{n-1}{2}\right),$$

or finally substituting for y_0 from (viii):

$$y = \frac{Nn^{\frac{n-1}{2}}}{2^{\frac{n-3}{2}}\Gamma\left(\frac{n-1}{2}\right)\Sigma_1} e^{-\frac{1}{2}\frac{n\sigma_1^2}{\Sigma_1^2}} \left(\frac{\sigma_1}{\Sigma_1}\right)^{n-2} \quad \text{.....(xi).}$$

This can be checked by integrating with regard to σ_1 and gives at once $\int_0^\infty y d\sigma_1 = N$, as it should. It is of course the usual form for the distribution of standard deviations in samples of n from a normal univariate population*. We now proceed to find

$$\int_0^\infty y\sigma_1 d\sigma_1 \text{ and } \int_0^\infty y\sigma_1^2 d\sigma_1$$

or as we may write them:

$$N_{\sigma_1}\mu_1' \text{ and } N_{\sigma_1}\mu_2'.$$

We have $$N_{\sigma_1}\mu_1' = y_0\Sigma_1^2 \int_0^\infty e^{-\frac{1}{2}\frac{n\sigma_1^2}{\Sigma_1^2}} \left(\frac{\sigma_1}{\Sigma_1}\right)^{n-1} d\left(\frac{\sigma_1}{\Sigma_1}\right),$$

where y_0 is written for brevity for

$$Nn^{\frac{n-1}{2}} \Big/ 2^{\frac{n-3}{2}}\Gamma\left(\frac{n-1}{2}\right)\Sigma_1.$$

Hence $$_{\sigma_1}\mu_1' = \frac{y_0}{N}\frac{\Sigma_1^2 2^{\frac{n-2}{2}}}{n^{\frac{1}{2}n}}\Gamma\left(\frac{n}{2}\right),$$

or substituting for y_0

$$_{\sigma_1}\mu_1' = \Sigma_1\sqrt{\frac{2}{n}}\frac{\Gamma\left(\frac{n}{2}\right)}{\Gamma\left(\frac{n-1}{2}\right)} \quad \text{.....(xii)}$$

Again $$N_{\sigma_1}\mu_2' = y_0\Sigma_1^3 \int_0^\infty e^{-\frac{1}{2}\frac{n\sigma_1^2}{\Sigma_1^2}} \left(\frac{\sigma_1}{\Sigma_1}\right)^n d\left(\frac{\sigma_1}{\Sigma_1}\right)$$

$$= \frac{y_0\Sigma_1^3 2^{\frac{n-1}{2}}}{n^{\frac{n-1}{2}}}\Gamma\left(\frac{n+1}{2}\right),$$

or, on substituting for y_0, we find the familiar result

$$_{\sigma_1}\mu_2' = \Sigma_1^2 . \frac{n-1}{n} \quad \text{.....(xiii).}$$

Hence writing $\lambda_n = \Gamma\left(\frac{n}{2}\right)\Big/\Gamma\left(\frac{n-1}{2}\right)$, we have

$$_{\sigma_1}\mu_2 = \sigma^2_{\sigma_1} = \frac{\Sigma_1^2}{n}(n - 1 - 2\lambda_n^2) \quad \text{.....(xiv).}$$

* *Biometrika*, Vol. VI. p. 304 and Vol. X. p. 522.

Accordingly

$$\left.\begin{aligned} \bar{x}_{\sigma_1}\bar{x}_{\sigma_2} &= \Sigma_1\Sigma_2\frac{2}{n}\lambda_n^2 \\ \sigma_{\sigma_1}\sigma_{\sigma_2} &= \Sigma_1\Sigma_2\frac{1}{n}(n-1-2\lambda_n^2) \end{aligned}\right\} \quad \text{.....(xv).}$$

We can write (x) in the form:

$$p'_{\sigma_1\sigma_2} = \Sigma_1\Sigma_2\frac{2}{n}\lambda_n^2 H(\rho^2),$$

and as the correlation coefficient:

$$r_{\sigma_1\sigma_2} = \frac{p'_{\sigma_1\sigma_2} - \bar{x}_{\sigma_1}\bar{x}_{\sigma_2}}{\sigma_{\sigma_1}\sigma_{\sigma_2}},$$

we find

$$r_{\sigma_1\sigma_2} = \frac{2\lambda_n^2(H(\rho^2)-1)}{n-1-2\lambda_n^2} \quad \text{.....(xvi).}$$

We shall now proceed to investigate the value of the factor $2\lambda_n^2/(n-1-2\lambda_n^2)$ when $n-1$ is *large*. We shall have to be cautious because if $2\lambda_n^2$ starts with the term $(n-1)$ everything will depend on the *second* term in the expansion. We will therefore expand λ_n in terms of $\epsilon = \frac{1}{2}(n-1)$.

A good formula for $\Gamma(x)$ is that provided by Schlömilch*, namely:

$$\log\Gamma(x) = \tfrac{1}{2}\log 2\pi + (x-\tfrac{1}{2})\log x - x + \frac{1}{12}\frac{1}{x} - \frac{1}{360}\frac{1}{x(x+1)(x+2)} + \text{etc.} \quad \text{(xvii).}$$

Now
$$\begin{aligned} \log\lambda_n &= \log\Gamma\left(\frac{n}{2}\right) - \log\Gamma\left(\frac{n-1}{2}\right) \\ &= \log\Gamma\left(\frac{n-1}{2}+\tfrac{1}{2}\right) - \log\Gamma\left(\frac{n-1}{2}\right) \\ &= \log\Gamma(\epsilon+\tfrac{1}{2}) - \log\Gamma(\epsilon), \end{aligned}$$

and this equals, by the above formula:

$$\begin{aligned} &\epsilon\log(\epsilon+\tfrac{1}{2}) - \epsilon - \tfrac{1}{2} + \frac{1}{12}\frac{1}{\epsilon+\frac{1}{2}} + \text{etc.} \\ &-(\epsilon-\tfrac{1}{2})\log\epsilon + \epsilon - \frac{1}{12}\frac{1}{\epsilon} + \text{etc.} \end{aligned}$$

or
$$\begin{aligned} \log\lambda_n &= \tfrac{1}{2}\log\epsilon + \epsilon\log\left(1+\frac{1}{2\epsilon}\right) - \tfrac{1}{2} + \text{terms in } \frac{1}{\epsilon^2} \text{ etc.} \\ &= \tfrac{1}{2}\log\epsilon + \tfrac{1}{2} - \frac{1}{8\epsilon} - \tfrac{1}{2} + \text{terms in } \frac{1}{\epsilon^2} \text{ etc.} \end{aligned}$$

Hence
$$\begin{aligned} \lambda_n &= \sqrt{\epsilon}\, e^{-\frac{1}{8\epsilon}} = \sqrt{\epsilon}\left(1-\frac{1}{8\epsilon}\right) + \text{etc.} \\ &= \frac{\sqrt{\epsilon}}{\left(1+\frac{1}{4\epsilon}\right)^{\frac{1}{2}}} \end{aligned}$$

to our degree of approximation.

* *Compendium der höheren Analysis*, Bd. II. S. 264.

Hence $$\lambda_n^2 = \frac{4\epsilon^2}{1+4\epsilon},$$

and $$\frac{2\lambda_n^2}{2\epsilon - 2\lambda_n^2} = \frac{\lambda_n^2}{\epsilon - \lambda_n^2} = \frac{4\epsilon^2}{\epsilon} = 2(n-1) + \text{etc.}$$

Accordingly when n is large:

$$r_{\sigma_1\sigma_2} \text{ approaches } (2n-2)(H(\rho^2)-1)$$
$$= \rho^2 + \text{terms in inverse powers of } 2n+2,\ 2n+4, \text{ etc.}$$

This approaches the value ρ^2 given long ago* for *large* samples from material with linear regression and homoscedasticity.

In the ordinary theory of probable errors it is usually the mean value of the product $p_{\sigma_1\sigma_2}$ that is involved, i.e. $\sigma_{\sigma_1}\sigma_{\sigma_2} r_{\sigma_1\sigma_2}$. But σ_{σ_1}, σ_{σ_2} are known and have been tabled for small samples. It remains to table $r_{\sigma_1\sigma_2}$ and see how far it diverges from the value usually assumed (i.e. ρ^2) for it.

The following table gives the values of $r_{\sigma_1\sigma_2}$ in samples of size 2, 3, ... 25, 50, 100, 400 for correlations ρ in the sampled population of ·0, ·1, ·2, ... ·8, ·9.

TABLE I. *Values of the Correlation of Standard Deviations in Samples for Different Values of Variate Correlation in a Sampled Normal Population.*

Values of ρ.

Size of Sample.	·0	·1	·2	·3	·4	·5	·6	·7	·8	·9	1·0
2	·0000	·0088	·0352	·0794	·1420	·2236	·3249	·4470	·5912	·7587	·9514
3	·0000	·0092	·0367	·0828	·1484	·2323	·3369	·4622	·6094	·7795⁻	·9737
4	·0000	·0094	·0375⁺	·0846	·1509	·2369	·3430	·4699	·6184	·7893	·9837
5	·0000	·0095⁻	·0380	·0857	·1528	·2396	·3467	·4749	·6235⁺	·7948	·9890
6	·0000	·0096	·0383	·0864	·1540	·2415⁻	·3491	·4775⁺	·6268	·7981	·9921
7	·0000	·0096	·0386	·0870	·1549	·2427	·3508	·4794	·6291	·8004	·9940
8	·0000	·0097	·0388	·0873	·1556	·2437	·3520	·4808	·6307	·8020	·9953
9	·0000	·0097	·0389	·0876	·1561	·2444	·3530	·4820	·6319	·8031	·9962
10	·0000	·0097	·0390	·0879	·1565⁻	·2450	·3537	·4829	·6328	·8040	·9969
11	·0000	·0098	·0391	·0881	·1568	·2455⁻	·3543	·4836	·6336	·8047	·9975⁻
12	·0000	·0098	·0392	·0882	·1571	·2459	·3548	·4842	·6342	·8053	·9978
13	·0000	·0098	·0392	·0884	·1573	·2462	·3552	·4847	·6347	·8058	·9982
14	·0000	·0098	·0393	·0885	·1575⁺	·2465⁻	·3556	·4851	·6351	·8061	·9984
15	·0000	·0098	·0393	·0886	·1577	·2467	·3559	·4854	·6355⁺	·8065⁻	·9986
16	·0000	·0098	·0394	·0887	·1578	·2469	·3562	·4857	·6358	·8067	·9988
17	·0000	·0099	·0394	·0888	·1580	·2471	·3564	·4860	·6361	·8070	·9989
18	·0000	·0099	·0395⁻	·0888	·1581	·2473	·3566	·4862	·6366	·8072	·9990
19	·0000	·0099	·0395⁻	·0889	·1582	·2474	·3568	·4864	·6363	·8074	·9991
20	·0000	·0099	·0395⁺	·0890	·1583	·2476	·3570	·4866	·6367	·8075⁺	·9992
21	·0000	·0099	·0395⁺	·0890	·1584	·2477	·3571	·4868	·6369	·8077	·9993
22	·0000	·0099	·0396	·0891	·1584	·2478	·3573	·4870	·6371	·8078	·9993
23	·0000	·0099	·0396	·0891	·1585⁺	·2479	·3574	·4871	·6372	·8079	·9994
24	·0000	·0099	·0396	·0891	·1586	·2480	·3575⁻	·4872	·6373	·8080	·9994
25	·0000	·0099	·0396	·0892	·1586	·2481	·3576	·4873	·6375⁻	·8081	·9995
50	·0000	·0100	·0398	·0896	·1593	·2490	·3588	·4887	·6388	·8091	·9999
100	·0000	·0100	·0399	·0898	·1597	·2495⁺	·3594	·4894	·6394	·8095⁺	·99997
400	·0000	·0100	·0400	·0900	·1599	·2499	·3599	·4898	·6399	·8099	1·00000
ρ^2	·0000	·0100	·0400	·0900	·1600	·2500	·3600	·4900	·6400	·8100	1·0000

* *Biometrika*, Vol. IX. p. 4. See also *Phil. Trans.* Vol. 191, A (1898), p. 242.

II. *On the Correlation of Deviations in the Standard Deviation of a Variate and the Correlation Coefficient of Variates in Samples from a Normal Population.*

We start from the general surface for the three variates, i.e.:

$$z = z_0' e^{-\frac{1}{2}\left(\frac{\sigma_1}{s_1}\right)^2} e^{-\frac{1}{2}\left(\frac{\sigma_2}{s_2}\right)^2} e^{\gamma r}(1-r^2)^{\frac{n-4}{2}} \left(\frac{\sigma_1}{s_1}\right)^{n-2}\left(\frac{\sigma_2}{s_2}\right)^{n-2},$$

and integrate it for the variate σ_2 between the limits 0 and ∞. We shall thus reach the bivariate frequency surface for σ_1 and r. In order to do this we must expand the exponential $e^{\gamma r}$ and find:

$$z' = \int_0^\infty z\, d\sigma_2 = z_0' e^{-\frac{1}{2}\left(\frac{\sigma_1}{s_1}\right)^2}\left(\frac{\sigma_1}{s_1}\right)^{n-2}(1-r^2)^{\frac{n-4}{2}} \int_0^\infty e^{-\frac{1}{2}\left(\frac{\sigma_2}{s_2}\right)^2}\left(\frac{\sigma_2}{s_2}\right)^{n-2} \sum_0^\infty \frac{(r\rho)^p}{p!}\left(\frac{\sigma_1\sigma_2}{s_1 s_2}\right)^p d\sigma_2.$$

Remembering that the integral

$$\int_0^\infty e^{-\frac{1}{2}\left(\frac{\sigma_2}{s_2}\right)^2}\left(\frac{\sigma_2}{s_2}\right)^{n-2+p} d\sigma_2 = s_2 2^{\frac{n-3+p}{2}}\Gamma\left(\frac{n-1+p}{2}\right),$$

we can write for the frequency surface of σ_1 and r

$$z' = z_0' s_2 e^{-\frac{1}{2}\left(\frac{\sigma_1}{s_1}\right)^2}\left(\frac{\sigma_1}{s_1}\right)^{n-2}(1-r^2)^{\frac{n-4}{2}} 2^{\frac{n-3}{2}}\left[\Gamma\left(\frac{n-1}{2}\right)\right.$$

$$\times\left\{1 + 2\frac{n-1}{2}\frac{(r\rho)^2}{2!}\left(\frac{\sigma_1}{s_1}\right)^2 + 2^2\frac{n-1}{2}\frac{n+1}{2}\frac{(r\rho)^4}{4!}\left(\frac{\sigma_1}{s_1}\right)^4 + \ldots\right\}$$

$$\left. + \Gamma\left(\frac{n}{2}\right)\sqrt{2}\, r\rho\,\frac{\sigma_1}{s_1}\left\{1 + 2\frac{n}{2}\frac{(r\rho)^2}{3!}\left(\frac{\sigma_1}{s_1}\right)^2 + 2^2\frac{n}{2}\left(\frac{n}{2}+1\right)\frac{(r\rho)^4}{5!}\left(\frac{\sigma_1}{s_1}\right)^4 + \ldots\right\}\right] \ldots(\text{xviii}).$$

The complexity of this surface for r and σ_1 is at once obvious, and it is not easy to form a mental picture of the form of the surface for small values of n. On the other hand we shall find it relatively easy to evaluate the integrals

$$I_{00} = \int_{-1}^{+1}\int_0^\infty z'\, dr\, d\sigma_1, \quad I_{10} = \int_{-1}^{+1}\int_0^\infty z' r\, dr\, d\sigma_1, \quad I_{01} = \int_{-1}^{+1}\int_0^\infty z'\sigma_1\, dr\, d\sigma_1,$$

$$I_{20} = \int_{-1}^{+1}\int_0^\infty z' r^2\, dr\, d\sigma_1, \text{ and } I_{11} = \int_{-1}^{+1}\int_0^\infty z' r\sigma_1\, dr\, d\sigma_1.$$

The first four will check results already known, and the last will provide us with what we are at present seeking. When we integrate with regard to r, all the terms in odd powers of r disappear and we may take for the even powers the double integral from 0 to +1. Remembering that

$$2\int_0^1 (1-r^2)^{\frac{n-4}{2}} r^{2p}\, dr = B\left(\frac{2p+1}{2},\ \frac{n-2}{2}\right)$$

$$= \Gamma\left(\frac{2p+1}{2}\right)\Gamma\left(\frac{n-2}{2}\right)\Big/\Gamma\left(\frac{2p+n-1}{2}\right)$$

we easily find

$$\int_{-1}^{+1} z' dr = z_0' s_2 e^{-\frac{1}{2}\left(\frac{\sigma_1}{s_1}\right)^2} \left(\frac{\sigma_1}{s_1}\right)^{n-2} 2^{\frac{n-3}{2}} \sqrt{\pi}\, \Gamma\left(\frac{n-2}{2}\right) \times \left\{1 + \frac{\rho^2}{2}\left(\frac{\sigma_1}{s_1}\right)^2 \frac{1}{1!} + \frac{\rho^4}{2^2}\left(\frac{\sigma_1}{s_1}\right)^4 \frac{1}{2!} + \dots\right\}$$

$$= z_0' s_2 e^{-\frac{1}{2}\left(\frac{\sigma_1}{s_1}\right)^2} e^{+\frac{1}{2}\rho^2 \frac{\sigma_1^2}{s_1^2}} 2^{\frac{n-3}{2}} \sqrt{\pi}\, \Gamma\left(\frac{n-2}{2}\right) \left(\frac{\sigma_1}{s_1}\right)^{n-2}$$

$$= z_0' s_2 e^{-\frac{1}{2}(1-\rho^2)\frac{\sigma_1^2}{s_1^2}} \left(\frac{\sigma_1}{s_1}\right)^{n-2} \sqrt{\pi}\, 2^{\frac{n-3}{2}} \Gamma\left(\frac{n-2}{2}\right) \dots\dots\dots\dots\dots\dots\text{(xix)}.$$

If we integrate next with regard to σ_1 we have to find

$$\int_0^\infty e^{-\frac{1}{2}(1-\rho^2)\frac{\sigma_1^2}{s_1^2}} \left(\frac{\sigma_1}{s_1}\right)^{n-2} d\sigma_1,$$

which equals

$$s_1 2^{\frac{n-3}{2}} \frac{1}{(1-\rho^2)^{\frac{n-1}{2}}} \Gamma\left(\frac{n-1}{2}\right).$$

Accordingly we have

$$N = z_0' s_1 s_2 \sqrt{\pi}\, 2^{n-3} \Gamma\left(\frac{n-2}{2}\right) \Gamma\left(\frac{n-1}{2}\right) \Big/ (1-\rho^2)^{\frac{n-1}{2}} \quad \dots\dots\dots\text{(xx)},$$

which checks with earlier results.

We will now investigate $\int_{-1}^{+1} z' r\, dr$, which throws us onto our second series in (xviii) for z' :

$$\int_{-1}^{+1} z' r\, dr = z_0' s_2 e^{-\frac{1}{2}\left(\frac{\sigma_1}{s_1}\right)^2} \left(\frac{\sigma_1}{s_1}\right)^{n-1} 2^{\frac{n-2}{2}} \rho \Gamma\left(\frac{n}{2}\right) \tfrac{1}{2} \sqrt{\pi} \frac{\Gamma\left(\frac{n-2}{2}\right)}{\Gamma\left(\frac{n+1}{2}\right)}$$

$$\times \left\{1 + \frac{n}{n+1}\frac{\rho^2}{1!\,2}\left(\frac{\sigma_1}{s_1}\right)^2 + \frac{n(n+2)}{(n+1)(n+3)}\frac{\rho^4}{2!\,2^2}\left(\frac{\sigma_1}{s_1}\right)^4 + \dots\right\} \quad \dots\text{(xxi)}.$$

If we divide this result by that in (xx) we obtain the mean value of r for a given array of r's corresponding to σ_1, i.e.

$$\bar{r}_{\sigma_1} = \frac{\rho}{\sqrt{2}} \frac{\Gamma\left(\frac{n}{2}\right)}{\Gamma\left(\frac{n+1}{2}\right)} \frac{\sigma_1}{s_1} e^{-\frac{1}{2}\rho^2\left(\frac{\sigma_1}{s_1}\right)^2} \left\{1 + \frac{n}{n+1}\frac{\rho^2}{1!\,2}\left(\frac{\sigma_1}{s_1}\right)^2 + \frac{n(n+2)}{(n+1)(n+3)}\frac{\rho^4}{2!\,2^2}\left(\frac{\sigma_1}{s_1}\right)^4 + \dots\right\} \dots\dots\dots\text{(xxii)}.$$

The regression line is therefore non-linear. Let us see what form it takes as we increase n. Since $s_1 = \Sigma_1 \sqrt{1-\rho^2}/\sqrt{n}$, we must approach the limit with caution

for s_1 becomes zero as n increases. Let us write $\kappa = \frac{1}{2}\rho^2\sigma_1^2/((1-\rho^2)\,\Sigma_1^2)$ which will be finite. Then we have to find the limit of

$$\bar{r}_{\sigma_1} = \rho\,\frac{\sqrt{\frac{1}{2}n}\,\Gamma\left(\frac{n}{2}\right)}{\Gamma(\frac{1}{2}(n+1))}\,\frac{\sigma_1}{\Sigma_1\sqrt{1-\rho^2}}\,e^{-n\kappa}\left(1+\frac{n}{n+1}\,\frac{n\kappa}{1!}+\frac{n\,(n+2)}{(n+1)(n+3)}\,\frac{n^2\kappa^2}{2!}\right.$$
$$\left.+\frac{n(n+2)\,(n+4)}{(n+1)\,(n+3)\,(n+5)}\,\frac{n^3\kappa^3}{3!}+\ldots\right).$$

But $\sqrt{\frac{1}{2}n}\,\Gamma\,(\frac{1}{2}n)/\Gamma\,(\frac{1}{2}\,(n+1))$ takes the limit unity.

Writing $n\kappa = \kappa'$, we have now to find the limit of:

$$L = e^{-\kappa'}\left(1+\frac{n}{n+1}\,\frac{\kappa'}{1!}+\frac{n\,(n+2)}{(n+1)(n+3)}\,\frac{\kappa'^2}{2!}+\frac{n\,(n+2)\,(n+4)}{(n+1)\,(n+3)\,(n+5)}\,\frac{\kappa'^3}{3!}+\ldots\right).$$

Expanding the exponential and multiplying out we have:

$$L = 1-\frac{1}{n+1}\,\frac{\kappa'}{1!}+\frac{1.3}{(n+1)\,(n+3)}\,\frac{\kappa'^2}{2!}-\frac{1.3.5}{(n+1)\,(n+3)\,(n+5)}\,\frac{\kappa'^3}{3!}+\ldots$$
$$= 1-\frac{n}{n+1}\,\frac{\kappa}{1!}+\frac{n^2\,1.3}{(n+1)\,(n+3)}\,\frac{\kappa^2}{2!}-\frac{n^3\,1.3.5}{(n+1)\,(n+3)\,(n+5)}\,\frac{\kappa^3}{3!}+\ldots.$$

Hence proceeding to the limit we find:

$$L = 1-\frac{\kappa}{1!}+\frac{1.3}{2!}\,\kappa^2-\frac{1.3.5}{3!}\,\kappa^3+\ldots$$
$$= (1+2\kappa)^{-\frac{1}{2}} = \left(1+\frac{\rho^2}{1-\rho^2}\,\frac{\sigma_1^2}{\Sigma_1^2}\right)^{-\frac{1}{2}}.$$

Or as n grows large we have nearly and more nearly:

$$\bar{r}_{\sigma_1} = \rho\,\frac{\sigma_1}{\Sigma_1}\,\frac{1}{\left(1-\rho^2+\rho^2\left(\frac{\sigma_1}{\Sigma_1}\right)^2\right)^{\frac{1}{2}}} \quad \ldots\ldots\ldots\ldots\ldots\ldots\text{(xxii } bis\text{)}.$$

Thus the equation suggested by normal correlation:

$$\bar{r}_{\sigma_1}-\bar{r} = \frac{\rho}{\sqrt{2}}\,\frac{\sigma_r}{\sigma_\sigma}\,(\sigma_1-\bar{\sigma}_1) = \rho\,.\,(1-\rho^2)\,\frac{\sigma_1-\bar{\sigma}_1}{\Sigma_1},$$

i.e. the usual value for linear regression, is seen to be incorrect. The true regression is not linear, even for n indefinitely large. When $\rho = 1$, $\bar{r}_{\sigma_1} = 1$, from (xxii *bis*), as it should. This result seems of considerable importance. For, while the distribution of σ_1 approaches closely to a normal curve when n is large*, and the values of β_1, β_2 for the distribution of r approach the values 0 and 3, i.e. the normal values, for moderate values of ρ and values over 50 of n, they do not do so for ρ above 0·9. In other words the limit for n large in the case of (xviii) is not a normal surface. The appearance of such a regression curve as (xxii *bis*) seems to me of the highest importance, and provides a warning against the too ready use of linear regression for frequency constants.

* *Biometrika*, Vol. x. p. 524.

If we now integrate the expression in (xxi) for σ_1 we obtain

$$N\bar{r} = z_0 s_1 s_2 2^{n-3} \rho \sqrt{\pi} \frac{\Gamma^2\left(\frac{n}{2}\right)\Gamma\left(\frac{n-2}{2}\right)}{\Gamma\left(\frac{n+1}{2}\right)}$$

$$\times \left(1 + \frac{\frac{1}{2}n \cdot \frac{1}{2}n}{\frac{1}{2}(n+1)} \frac{\rho^2}{1!} + \frac{\frac{1}{2}n\,(\frac{1}{2}n+1)\,\frac{1}{2}n\,(\frac{1}{2}n+1)}{\frac{1}{2}(n+1)\,\frac{1}{2}(n+3)} \frac{\rho^4}{2!} + \ldots\right).$$

But the series $= F(\frac{1}{2}n, \frac{1}{2}n, \frac{1}{2}(n+1), \rho^2)$,

where $F(\alpha, \beta, \gamma, x)$ is the usual hypergeometrical series. Using Euler's transformation, we find:

$$N\bar{r} = \frac{z_0 s_1 s_2 2^{n-3} \rho \sqrt{\pi}}{(1-\rho^2)^{\frac{n-1}{2}}} \frac{\Gamma^2\left(\frac{n}{2}\right)\Gamma\left(\frac{n-2}{2}\right)}{\Gamma\left(\frac{n+1}{2}\right)} F(\tfrac{1}{2}, \tfrac{1}{2}, \tfrac{1}{2}(n+1), \rho^2).$$

Whence dividing by N from (xx)

$$\bar{r} = \rho \frac{\Gamma^2\left(\frac{n}{2}\right)}{\Gamma\left(\frac{n-1}{2}\right)\Gamma\left(\frac{n+1}{2}\right)} F(\tfrac{1}{2}, \tfrac{1}{2}, \tfrac{1}{2}(n+1), \rho^2) \ldots\ldots\ldots\ldots\text{(xxiii)}.$$

This agrees with *Biometrika*, Vol. XI. p. 336, Equation (xxv), if we change the notation and observe that the symbol q_n there used is the same as $B(\frac{1}{2}, \frac{1}{2}n)$.

It is extremely easy in the same manner to find $N\bar{\sigma}_1$. All we have to do is to integrate the expression in (xix) for $\sigma_1 d\sigma_1$, i.e.

$$N\bar{\sigma}_1 = \int_{-1}^{+1}\int_0^\infty z' \sigma_1 \, dr \, d\sigma_1 = z_0 s_1^2 s_2 \sqrt{\pi}\, 2^{\frac{n-3}{2}} \Gamma\left(\frac{n-2}{2}\right) \times \int_0^\infty e^{-\frac{1}{2}(1-\rho^2)\left(\frac{\sigma_1}{s_1}\right)^2} \left(\frac{\sigma_1}{s_1}\right)^{n-1} d\left(\frac{\sigma_1}{s_1}\right)$$

$$= \frac{z_0 s_1^2 s_2 \sqrt{\pi}\, 2^{\frac{2n-5}{2}}}{(1-\rho^2)^{\frac{1}{2}n}} \Gamma\left(\frac{n-2}{2}\right)\Gamma\left(\frac{n}{2}\right).$$

Whence getting rid of z_0 we find:

$$\bar{\sigma}_1 = \frac{s_1\sqrt{2}}{\sqrt{1-\rho^2}} \frac{\Gamma\left(\frac{n}{2}\right)}{\Gamma\left(\frac{n-1}{2}\right)} = \Sigma_1 \sqrt{\frac{2}{n}} \frac{\Gamma\left(\frac{n}{2}\right)}{\Gamma\left(\frac{n-1}{2}\right)} \ldots\ldots\ldots\text{(xxiv)}.$$

This agrees with the result (xii) given on p. 180.

We now proceed to find:

$$Np'_{r\sigma_1} = \int_0^\infty \sigma_1 \, d\sigma_1 \int_{-1}^{+1} z' r \, dr.$$

We return to Equation (xxi) and integrate with regard to $\sigma_1 d\sigma_1$:

$$Np'_{r\sigma_1} = z_0 s_2 s_1^2\, 2^{\frac{n-4}{2}} \rho \sqrt{\pi} \frac{\Gamma\left(\frac{n-2}{2}\right)}{\Gamma\left(\frac{n+1}{2}\right)} \Gamma\left(\frac{n}{2}\right) \times \int_0^\infty \left\{1 + \frac{n}{n+1} \frac{\rho^2}{1!\,2} \left(\frac{\sigma_1}{s_1}\right)^2\right.$$

$$+\frac{n(n+2)}{(n+1)(n+3)}\frac{\rho^4}{2!\,2^2}\left(\frac{\sigma_1}{s_1}\right)^4+\ldots\Big\}e^{-\frac{1}{2}\left(\frac{\sigma_1}{s_1}\right)^2}\left(\frac{\sigma_1}{s_1}\right)^n d\left(\frac{\sigma_1}{s_1}\right)$$

$$=z_0s_2s_1^2\,2^{\frac{2n-5}{2}}\,\Gamma\left(\frac{n+1}{2}\right)\rho\sqrt{\pi}\,\frac{\Gamma\left(\frac{n-2}{2}\right)\Gamma\left(\frac{n}{2}\right)}{\Gamma\left(\frac{n+1}{2}\right)}$$

$$\times\left(1+2\frac{n+1}{2}\frac{n}{n+1}\frac{\rho^2}{1!\,2}+2^2\frac{n+1}{2}\frac{n+3}{2}\frac{n(n+2)}{(n+1)(n+3)}\frac{\rho^4}{2!\,2^2}+\ldots\right).$$

Or:
$$Np'_{r\sigma_1}=\frac{z_0s_2s_1^2\rho\sqrt{\pi}\,2^{\frac{2n-5}{2}}\,\Gamma\left(\frac{n-2}{2}\right)\Gamma\left(\frac{n}{2}\right)}{(1-\rho^2)^{\frac{n}{2}}}.$$

Hence getting rid of z_0, we find

$$p'_{r\sigma_1}=\frac{s_1\rho\sqrt{2}}{\sqrt{1-\rho^2}}\frac{\Gamma\left(\frac{n}{2}\right)}{\Gamma\left(\frac{n-1}{2}\right)}=\Sigma_1\rho\sqrt{\frac{2}{n}}\frac{\Gamma\left(\frac{2}{n}\right)}{\Gamma\left(\frac{n-1}{2}\right)}\quad\ldots\ldots\ldots\ldots\text{(xxv)}.$$

We have already given in this paper the value of $\sigma_{\sigma_1}=\Sigma_1(n-1-2\lambda_n^2)^{\frac{1}{2}}/\sqrt{n}$, see Equation (xiv). In *Biometrika*, Vol. XI. p. 335, the value of σ_r^2 is given, but it may be well to deduce it by the methods here used. Clearly we must integrate z' for r^2dr, or better for $(1-(1-r^2))\,dr$ and then for $d\sigma_1$. It is seen at once that only the first series on the right of Equation (xviii) concerns us for the second contains odd powers. Accordingly we have

$$N(\sigma_r^2+\bar{r}^2)=\int_{-1}^{+1}z'(1-(1-r^2))\,dr\,d\sigma_1$$

$$=N-2z_0s_2\int_0^1\int_0^\infty(1-r^2)^{\frac{n-2}{2}}\,2^{\frac{n-3}{2}}\,\Gamma\left(\frac{n-1}{2}\right)e^{-\frac{1}{2}\left(\frac{\sigma_1}{s_1}\right)^2}\left(\frac{\sigma_1}{s_1}\right)^{n-2}$$

$$\times\left\{1+\frac{(n-1)}{2!}r^2\rho^2\left(\frac{\sigma_1}{s_1}\right)^2+\frac{(n-1)(n+1)}{4!}r^4\rho^4\left(\frac{\sigma_1}{s_1}\right)^4+\ldots\right\}dr\,d\sigma_1.$$

We have now

$$2\int_0^1(1-r^2)^{\frac{n-2}{2}}r^{2p}dr=\Gamma\left(\frac{2p+1}{2}\right)\Gamma\left(\frac{n}{2}\right)\Big/\Gamma\left(\frac{2p+n+1}{2}\right),$$

and accordingly

$$N(\sigma_r^2+\bar{r}^2)=N-z_0s_2\,2^{\frac{n-3}{2}}\cdot\frac{\Gamma\left(\frac{n}{2}\right)\Gamma\left(\frac{n-1}{2}\right)\sqrt{\pi}}{\Gamma\left(\frac{n+1}{2}\right)}\int_0^\infty e^{-\frac{1}{2}\left(\frac{\sigma_1}{s_1}\right)^2}\left(\frac{\sigma_1}{s_1}\right)^{n-2}$$

$$\times\left\{1+\tfrac{1}{2}\frac{n-1}{2!}\rho^2\left(\frac{\sigma_1}{s_1}\right)^2\frac{1}{\frac{n+1}{2}}+\tfrac{1}{2}\cdot\tfrac{3}{2}\frac{(n-1)(n+1)}{4!}\rho^4\left(\frac{\sigma_1}{s_1}\right)^4\frac{1}{\frac{n+1}{2}\frac{n+3}{2}}+\ldots\right\}d\sigma_1.$$

We can now integrate with regard to σ_1 and we have:

$$N(\sigma_r^2+\bar{r}^2) = N - z_0 s_1 s_2 2^{n-3}\sqrt{\pi}\,\frac{\Gamma\left(\frac{n}{2}\right)\Gamma^2\left(\frac{n-1}{2}\right)}{\Gamma\left(\frac{n+1}{2}\right)}$$

$$\times\left(1+\frac{n-1}{2}\frac{n-1}{2}\frac{\rho^2}{1!}\frac{1}{\frac{n+1}{2}}+\frac{n-1}{2}\frac{n+1}{2}\frac{n-1}{2}\frac{n+1}{2}\frac{\rho^4}{2!}\frac{1}{\frac{n+1}{2}\frac{n+3}{2}}+\text{etc.}\right)$$

$$= N - z_0 s_1 s_2\sqrt{\pi}\,2^{n-3}\,\frac{\Gamma\left(\frac{n}{2}\right)\Gamma^2\left(\frac{n-1}{2}\right)}{\Gamma\left(\frac{n+1}{2}\right)}\,F\left(\frac{n-1}{2},\ \frac{n-1}{2},\ \frac{n+1}{2},\ \rho^2\right),$$

or by Euler's transformation:

$$\sigma_r^2+\bar{r}^2 = 1-\frac{z_0 s_1 s_2\sqrt{\pi}\,2^{n-3}}{N(1-\rho^2)^{\frac{n-3}{2}}}\,\frac{\Gamma\left(\frac{n}{2}\right)\Gamma^2\left(\frac{n-1}{2}\right)}{\Gamma\left(\frac{n+1}{2}\right)}F(1,\ 1,\ \tfrac{1}{2}(n+1),\ \rho^2).$$

Substituting for N from (xx), we find:

$$\sigma_r^2 = 1-\bar{r}^2-(1-\rho^2)\,\frac{\Gamma\left(\frac{n}{2}\right)\Gamma\left(\frac{n-1}{2}\right)}{\Gamma\left(\frac{n-2}{2}\right)\Gamma\left(\frac{n+1}{2}\right)}F(1,\ 1,\ \tfrac{1}{2}(n+1),\ \rho^2)$$

$$= 1-\bar{r}^2-(1-\rho^2)\,\frac{n-2}{n-1}F(1,\ 1,\ \tfrac{1}{2}(n+1),\ \rho^2)\ \ldots\ldots\ldots\ldots\ldots\ldots\text{(xxvi)}.$$

With a different notation this accords with the result in *Biometrika*, Vol. XI. p. 335, Equation (xx) *bis*.

We can now write down the value of $\bar{p}_{r\sigma_1} = p'_{r\sigma_1} - \bar{r}\bar{\sigma}_1$ from Equations (xxiii)—(xxv) and we find finally, if

$$\lambda_n = \Gamma\left(\frac{n}{2}\right)\Big/\,\Gamma\left(\frac{n-1}{2}\right),$$

that
$$r_{r\sigma_1} = \rho\sqrt{\frac{2\lambda_n^2}{n-1-2\lambda_n^2}}\,\frac{\left(1-\frac{\lambda_n}{\lambda_n+1}F\left(\frac{1}{2},\ \frac{1}{2},\ \frac{n+1}{2},\ \rho^2\right)\right)}{\sigma_r}\ \ldots\text{(xxvi } \textit{bis}).$$

Noting (xxiii), this reduces to:

$$r_{r\sigma_1} = \sqrt{\frac{2\lambda_n^2}{n-1-2\lambda_n^2}}\,\frac{(\rho-\bar{r})}{\sigma_r}\ \ldots\ldots\ldots\ldots\ldots\ldots\text{(xxvii)},$$

and σ_r must be found from (xxv).

As the values of $\bar{r}$ and σ_r have been already found* and the radical has been computed for our earlier work in this paper, there is not much difficulty in obtaining $r_{r\sigma_1}$. The following table gives the numerical values, and shows the degree of approximation to the usual value $\rho/\sqrt{2}$.

* *Biometrika*, Vol. XI. pp. 372, 379—401.

TABLE II. *Values of the Correlation of the Standard Deviation of One Variate with the Correlation of Both Variates in Samples of Different Sizes taken from a Normal Population.*

Values of ρ.

Size of Sample.	·0	·1	·2	·3	·4	·5	·6	·7	·8	·9	1·0
2	·0000	·0481	·0958	·1431	·1893	·2340	·2762	·3141	·3438	·3532	·0000
3	·0000	·0581	·1156	·1726	·2284	·2822	·3328	·3779	·4122	·4188	·0000
4	·0000	·0620	·1237	·1847	·2445$^+$	·3026	·3575$^+$	·4073	·4469	·4608	·0000
5	·0000	·0641	·1279	·1911	·2533	·3138	·3717	·4250$^-$	·4696	·4931	·0000
6	·0000	·0654	·1305$^+$	·1952	·2586	·3210	·3809	·4369	·4857	·5179	·4134
7	·0000	·0663	·1323	·1979	·2626	·3260	·3874	·4455$^-$	·4876	·5368	·5069
8	·0000	·0669	·1336	·1998	·2653	·3297	·3922	·4519	·5067	·5512	·5550$^-$
9	·0000	·0674	·1345$^+$	·2013	·2674	·3325$^-$	·3959	·4569	·5137	·5624	·5845$^+$
10	·0000	·0677	·1353	·2025$^-$	·2691	·3347	·3989	·4609	·5194	·5712	·6045$^+$
11	·0000	·0680	·1359	·2034	·2704	·3365$^-$	·4013	·4642	·5240	·5783	·6189
12	·0000	·0683	·1364	·2042	·2715$^-$	·3380	·4033	·4669	·5278	·5841	·6298
13	·0000	·0685$^-$	·1368	·2048	·2724	·3392	·4050$^-$	·4692	·5310	·5888	·6383
14	·0000	·0686	·1372	·2054	·2732	·3403	·4064	·4712	·5338	·5928	·6452
15	·0000	·0688	·1375$^-$	·2059	·2739	·3412	·4077	·4729	·5361	·5962	·6508
16	·0000	·0689	·1377	·2063	·2744	·3420	·4088	·4743	·5381	·5992	·6554
17	·0000	·0690	·1380	·2066	·2750$^-$	·3427	·4097	·4756	·5399	·6017	·6594
18	·0000	·0691	·1382	·2069	·2754	·3434	·4106	·4767	·5414	·6040	·6628
19	·0000	·0692	·1383	·2072	·2758	·3439	·4113	·4777	·5428	·6058	·6657
20	·0000	·0693	·1385$^+$	·2075$^-$	·2762	·3444	·4120	·4786	·5441	·6076	·6683
21	·0000	·0694	·1386	·2077	·2765$^-$	·3448	·4127	·4795$^-$	·5452	·6092	·6706
22	·0000	·0694	·1388	·2079	·2768	·3452	·4131	·4802	·5462	·6105$^+$	·6726
23	·0000	·0695$^-$	·1389	·2081	·2771	·3456	·4136	·4809	·5470	·6118	·6744
24	·0000	·0695$^+$	·1390	·2083	·2773	·3460	·4141	·4815$^-$	·5479	·6130	·6761
25	·0000	·0696	·1391	·2085$^-$	·2775$^+$	·3463	·4145$^-$	·4820	·5487	·6140	·6775$^+$
50	·0000	·0701	·1403	·2104	·2802	·3500$^+$	·4195$^+$	·4886	·5576	·6258	·6936
100	·0000	·0704	·1409	·2112	·2815$^+$	·3518	·4219	·4919	·5617	·6314	·7007
400	·0000	·0706	·1413	·2119	·2825$^+$	·3531	·4237	·4942	·5647	·6350$^+$	·7055
$\rho/\sqrt{2}$	·0000	·0707	·1414	·2121	·2828	·3536	·4243	·4950	·5657	·6364	·7071

It will be clear to the reader that this table possesses more singularities than that for $r_{\sigma_1\sigma_2}$. The maximum value for $r_{\sigma_1\sigma_2}$ for a sample of definite size n may be reached only at $\rho = 1{\cdot}0$, or it may in the case of samples of 8 or less be reached earlier. Further it may be doubted whether for practical purposes we can use the ordinary value, i.e. that for "large" samples ($\rho/\sqrt{2}$) at $n = 25$; certainly only roughly for $\rho = {\cdot}7$ and over. We might possibly for *practical* purposes fix the limit of small samples at $n = 50$. If we want, however, the probable error of such a simple constant as the regression coefficient, $R = r\sigma_2/\sigma_1$, for a small sample it is needful actually to know not only $\{\delta r \delta\sigma_1\}$, $\{\delta r \delta\sigma_2\}$, $\{\delta\sigma_1\delta\sigma_2\}$, but we clearly require higher products such as $\{\delta\sigma_1{}^2\}$, $\{\delta r \delta\sigma_1 \delta\sigma_2\}$, $\{\delta\sigma_1\delta\sigma_2{}^2\}$, etc. and also the fourth order products to obtain $\sigma_R{}^2$ to what is after all only then a second order approximation. If we go to the roughest form of approximation only, we have*:

$$\sigma_R{}^2 = \breve{R}^2\left(\frac{\sigma_r{}^2}{\breve{\rho}^2} + \frac{\sigma^2{}_{\sigma_1}}{\breve{\sigma}_1{}^2} + \frac{\sigma^2{}_{\sigma_2}}{\breve{\sigma}_2{}^2} - \frac{2\{\delta\sigma_1\delta\sigma_2\}}{\breve{\sigma}_1\breve{\sigma}_2} - \frac{2\{\delta\sigma_1\delta r\}}{\breve{\sigma}_1\breve{\rho}} + \frac{2\{\delta\sigma_2\delta r\}}{\breve{\sigma}_2\breve{\rho}}\right) \ldots(\text{xxviii}).$$

* A short quantity ~ over the letter indicates that the value should be that of the population sampled.

Now such an approximation will show us how far the use of large sample values for σ_r, σ_{σ_1}, σ_{σ_2}, $\{\delta\sigma_1\delta\sigma_2\}$, $\{\delta\sigma_1\delta r\}$, $\{\delta\sigma_2\delta r\}$ modifies the customary value of σ_R.

To illustrate this let us suppose $\rho = \cdot 6$ and the sample to be of size 10. Then the value of σ_R for a large sample would be $\check{R}\sqrt{1-\rho^2}/(\rho\sqrt{N})$*. This equals for our case $\cdot 4216\check{R}$.

We will now insert in the above formula the values of the several terms:

$$\frac{\sigma_r}{\check{\rho}} = \frac{\cdot 2355}{\cdot 6} = \cdot 3925\ \dagger,$$

$$\frac{\sigma_{\sigma_1}}{\sigma_1} = \cdot 2203^{+}_{+}, \quad \frac{\sigma_{\sigma_2}}{\sigma_2} = \cdot 2203^{+}_{+},$$

$$\{\delta\sigma_1\delta\sigma_2\} = \sigma_{\sigma_1}\sigma_{\sigma_2}r_{\sigma_1\sigma_2} = \check{\sigma}_1\check{\sigma}_2(\cdot 2203)^2\cdot 3537 = \check{\sigma}_1\check{\sigma}_2\cdot 0171{,}6580,$$

$$\{\delta\sigma_1\delta r\ \} = \sigma_{\sigma_1}\sigma_r\, r_{r\sigma_1} = \check{\rho}\,\check{\sigma}_1(\cdot 3925)(\cdot 2203)\cdot 3989 = \check{\rho}\check{\sigma}_1\cdot 0344{,}9199,$$

$$\{\delta\sigma_2\delta r\ \} = \sigma_{\sigma_2}\sigma_r\, r_{r\sigma_2} = \check{\rho}\,\check{\sigma}_2(\cdot 3925)(\cdot 2203)\cdot 3989 = \check{\rho}\check{\sigma}_2\cdot 0344{,}9199,$$

these last three values being deduced from the present paper. Accordingly the terms in $\{\delta\sigma_1\delta r\}$ and $\{\delta\sigma_2\delta r\}$ cancel and we have

$$\begin{aligned}\sigma_R^2 &= \check{R}^2(\cdot 1540{,}5625 + 2\times\cdot 0485{,}3209 - 2\times\cdot 0171{,}6580)\\ &= \check{R}^2\times\cdot 2167{,}8883,\end{aligned}$$

or $$\sigma_R = \cdot 4205\check{R}.$$

This is so close to $\cdot 4216\check{R}$ found for σ_R from the usual formula for large samples, that it is clear that grave error would not be introduced by using the ordinary values of $\sigma_r/\check{\rho}$ and σ_σ/σ, together with that for $\{\delta\sigma_1\delta\sigma_2\}$ in the case of $n = 10$. It is a different question, however, whether it is legitimate to neglect the contributions of the third and fourth order products to σ_r^2, when $1/N$ is of the order $1/10$. The answer to this question cannot however be provided until these higher products have been studied and we know their numerical values for small samples. We shall return to this point in our third section.

For the special case of $r_{r\sigma_1}$ when $\rho = 1$ we require to ascertain a limit. For in this case $\bar{r}$ will equal $\rho = 1$, and σ_r will be zero so that the ratio $(\rho - \bar{r})/\sigma_r$ takes the indeterminant form $0/0$. In order to evaluate the limit we require to expand $\bar{r}$ and σ_r in powers of ϵ where $\rho^2 = 1 - \epsilon$ and ϵ is made very small. First let us consider $\bar{r}$ in Equation (xxiii), we have

$$\bar{r} = \sqrt{1-\epsilon}\,\psi_n F(\tfrac{1}{2}, \tfrac{1}{2}, \tfrac{1}{2}(n+1), 1-\epsilon),$$

where $$\psi_n = \Gamma^2\left(\frac{n}{2}\right)\Big/\left\{\Gamma\left(\frac{n-1}{2}\right)\cdot\Gamma\left(\frac{n+1}{2}\right)\right\}.$$

Hence $$\bar{r} = (1 - \tfrac{1}{2}\epsilon - \tfrac{1}{8}\epsilon^2 + \ldots)\,\psi_n\left\{F_1 - \epsilon\left(\frac{dF}{dx}\right)_1 + \frac{\epsilon^2}{2}\left(\frac{d^2F}{dx^2}\right)_1 - \ldots\right\},$$

* *Phil. Trans.* Vol. 191 A, p. 245 or *Biometrika*, Vol. IX. p. 9.

† *Biometrika*, Vol. XI. p. 386.

‡ *Ibid.* Vol. X. p. 529.

where F_1 stands for $F(\frac{1}{2}, \frac{1}{2}, \frac{1}{2}(n+1), 1)$ and $\left(\frac{d^sF}{dx^s}\right)_1$ for the sth differential coefficient of $F(\frac{1}{2}, \frac{1}{2}, \frac{1}{2}(n+1), x)$, x being put unity after differentiation.

But it is well-known that

$$F(\alpha, \beta, \gamma, 1) = \frac{\Gamma(\gamma)\,\Gamma(\gamma-\alpha-\beta)}{\Gamma(\gamma-\alpha)\,\Gamma(\gamma-\beta)},$$

or we have

$$F_1(\tfrac{1}{2}, \tfrac{1}{2}, \tfrac{1}{2}(n+1), 1) = \frac{\Gamma(\frac{1}{2}(n-1))\,\Gamma(\frac{1}{2}(n+1))}{\Gamma^2\left(\frac{n}{2}\right)} = \frac{1}{\psi_n}.$$

Thus $$\bar{r} = (1 - \tfrac{1}{2}\epsilon - \tfrac{1}{8}\epsilon^2 + \ldots)\left\{1 - \frac{1}{F_1}\left(\frac{dF}{dx}\right)_1 \epsilon + \frac{1}{2F_1}\left(\frac{d^2F}{dx^2}\right)_1 \epsilon^2 - \ldots\right\}.$$

Further it is easy to show from the differential equation of the hypergeometrical series that:

$$\frac{1}{F_1}\left(\frac{dF}{dx}\right)_1 = \frac{\alpha\beta}{\gamma-\alpha-\beta-1}, = \frac{1}{2(n-3)} \text{ in our case,}$$

$$\frac{1}{F_1}\left(\frac{d^2F}{dx^2}\right)_1 = \frac{\alpha\beta(\alpha\beta+\alpha+\beta+1)}{(\gamma-\alpha-\beta-1)(\gamma-\alpha-\beta-2)}, = \frac{9}{4}\frac{1}{(n-3)(n-5)} \text{ in our case.}$$

Hence we find

$$\rho - \bar{r} = \frac{1}{2(x-3)}\epsilon - \frac{2n-1}{8(n-3)(n-5)}\epsilon^2 + \ldots,$$

and

$$\bar{r} = 1 - \tfrac{1}{2}\frac{n-2}{n-3}\epsilon - \frac{(n-2)(n-8)}{8(n-3)(n-5)}\epsilon^2 + \ldots.$$

Accordingly

$$1 - \bar{r}^2 = \frac{n-2}{n-3}\epsilon - \frac{(2n-7)(n-2)}{2(n-3)^2(n-5)}\epsilon^2 + \ldots.$$

We now require

$$(1-\rho^2)\frac{n-2}{n-1}F(1, 1, \tfrac{1}{2}(n+1), \rho^2),$$

or

$$\epsilon\frac{n-2}{n-1}\left\{F_2 - \epsilon\left(\frac{dF_2}{dx}\right)_1 + \tfrac{1}{2}\epsilon^2\left(\frac{d^2F_2}{dx^2}\right)_1 - \ldots\right\},$$

where $F_2 = F(1, 1, \frac{1}{2}(n+1), \rho^2)$ and $\left(\frac{d^sF_2}{dx^s}\right)_1$ = the sth differential coefficient of $F = F_2(1, 1, \frac{1}{2}(n+1), x)$ with x put unity after differentiation.

But $$F_2 = \Gamma\left(\frac{n+1}{2}\right)\Gamma\left(\frac{n-1}{2}\right)\Big/\Gamma^2\left(\frac{n-1}{2}\right) = \frac{n-1}{n-3},$$

and $$\frac{1}{F_2}\left(\frac{dF_2}{dx}\right)_1 = \frac{2}{n-5}.$$

Thus from (xxvi) we have

$$\sigma_r^2 = \epsilon^2\left\{\frac{2(n-2)}{(n-3)(n-5)} - \frac{(n-2)(2n-7)}{2(n-3)^2(n-5)}\right\} + \text{etc.}$$

$$= \tfrac{1}{2}\epsilon^2\frac{(n-2)(2n-5)}{(n-3)^2(n-5)}.$$

Hence from the values provided on p. 191:

$$\mathop{\mathrm{L}}_{\rho=1} \frac{\rho-\bar{r}}{\sigma_r} = \sqrt{\frac{(n-5)}{2(n-2)(2n-5)}} \quad \text{.....(xxix),}$$

or
$$r_{r\sigma_1} \text{ for } \rho=1 = \sqrt{\frac{2\lambda_n^2}{n-1-2\lambda_n^2}} \sqrt{\frac{(n-5)}{2(n-2)(2n-5)}} \quad \text{.....(xxx).}$$

When $n=5$ this limiting value is zero. For $n=6$ and higher values, it is always real and positive. For $n=2, 3$ and 4 this method of approaching the limit fails—the values are imaginary. These three cases require separate investigation.

In the case of $n=2$, $\bar{r} = \frac{2}{\pi}\sin^{-1}\rho$ * and if $\rho = 1-\lambda$, we have:

$$\sin^{-1}\rho = \frac{\pi}{2} - 2\sin^{-1}\sqrt{\frac{\lambda}{2}}.$$

Therefore
$$\bar{r} = 1 - \frac{4}{\pi}\left\{\left(\frac{\lambda}{2}\right)^{\frac{1}{2}} + \tfrac{1}{6}\left(\frac{\lambda}{2}\right)^{\frac{3}{2}} + \ldots\right\},$$

$$\rho - \bar{r} = \lambda + \frac{4}{\pi}\left\{\left(\frac{\lambda}{2}\right)^{\frac{1}{2}} + \tfrac{1}{6}\left(\frac{\lambda}{2}\right)^{\frac{3}{2}} + \ldots\right\},$$

$$\bar{r}^2 = 1 - \frac{8}{\pi}\left\{\left(\frac{\lambda}{2}\right)^{\frac{1}{2}} + \tfrac{1}{6}\left(\frac{\lambda}{2}\right)^{\frac{3}{2}} + \ldots\right\} + \frac{8}{\pi^2}\lambda.$$

But for $n=2$, $\sigma_r^2 = 1-\bar{r}^2$†, or

$$\sigma_r^2 = \frac{8}{\pi}\left(\frac{\lambda}{2}\right)^{\frac{1}{2}} - \frac{8}{\pi^2}\lambda + \ldots.$$

Accordingly
$$\frac{\rho-\bar{r}}{\sigma_r} = \sqrt{2\pi}\,(\tfrac{1}{2}\lambda)^{\frac{1}{2}} + \text{etc.} \quad \text{.....(xxxi),}$$

and this is zero when $\lambda = 0$. The important point is that $\bar{r}$ starts with powers of $\lambda^{\frac{1}{2}}$ and not with integer powers.

We next turn to the case of $n=3$. Here we have‡:

$$\left.\begin{aligned} \bar{r} &= \frac{1}{\rho}\left(E_1 - (1-\rho^2)F_1\right) \\ \sigma_r^2 &= 1 - \bar{r}^2 + \tfrac{1}{2}\cdot\frac{1-\rho^2}{\rho}\log_e(1-\rho^2) \end{aligned}\right\} \quad \text{.....(xxxii),}$$

where E_1 and F_1 are the complete elliptic integrals for $k^2 = \rho^2$. The difficulty of the limit here is that as ρ approaches unity E_1 approaches unity and F_1 approaches infinity. But without proceeding to a limit we have by the tables of the complete elliptic integrals:

$\rho = \cdot 95$, $\quad \bar{r} = \cdot 894{,}94275$§, $\quad \sigma_r = \cdot 270{,}7989$§,
$\rho = \cdot 999{,}8477$, $\quad \bar{r} = \cdot 999{,}2484$, $\quad \sigma_r = \cdot 016{,}41104$,
$\rho = \cdot 999{,}9985$, $\quad \bar{r} = \cdot 999{,}987{,}813$, $\quad \sigma_r = \cdot 002{,}30196$,

* *Biometrika*, Vol. XI. p. 360. † *Ibid.* p. 361. ‡ *Ibid.* p. 363.
§ Erroneous numerical values for these constants are given, *Biometrika*, Vol. XI. p. 368.

giving respectively

$$r_{r\sigma} = \cdot 38895, \ = \cdot 06986 \text{ and } = \cdot 01013,$$

showing the rapid approach to zero of the value of $r_{r\sigma}$ as ρ approaches unity.

We next turn to the case of $n = 4$.

Here if $\rho = \sin\alpha$, then

$$\left.\begin{aligned} \bar{r} &= \frac{2}{\pi}\{\cot\alpha + \alpha(1-\cot^2\alpha)\} \\ \text{and} \quad \sigma_{\bar{r}}^2 &= 1 - \bar{r}^2 - 2\cot^2\alpha + 2\alpha\cot^3\alpha \end{aligned}\right\} \quad \ldots\ldots\ldots\ldots \text{(xxxiii)}.$$

Now let $\alpha = \frac{\pi}{2} - \gamma$, where γ is a small quantity. Then

$$\rho = \cos\gamma = 1 - \frac{\gamma^2}{2} + \ldots,$$

$$\begin{aligned} \bar{r} &= \frac{2}{\pi}\left\{\tan\gamma + \left(\frac{\pi}{2} - \gamma\right)(1-\tan^2\gamma)\right\} \\ &= \frac{2}{\pi}\left\{\gamma + \tfrac{1}{3}\gamma^3 + \left(\frac{\pi}{2} - \gamma\right)(1 - \gamma^2 + \ldots)\right\} \\ &= \frac{2}{\pi}\left(\tfrac{1}{3}\gamma^3 + \frac{\pi}{2} - \frac{\pi}{2}\gamma^2 + \gamma^3\right) \\ &= 1 - \gamma^2 + \frac{8}{3\pi}\gamma^3 + \ldots. \end{aligned}$$

Accordingly $\quad 1 - \bar{r}^2 = 2\gamma^2 - \dfrac{16}{3\pi}\gamma^3 + \ldots,$

$$\rho - \bar{r} = \tfrac{1}{2}\gamma^2 - \text{terms in } \gamma^3,$$

$$\begin{aligned} \sigma_{\bar{r}}^2 &= 1 - \bar{r}^2 - 2\tan^2\gamma + 2\left(\frac{\pi}{2} - \gamma\right)\tan^3\gamma \\ &= \left(\pi - \frac{16}{3\pi}\right)\gamma^3 - \tfrac{7}{3}\gamma^4 + \ldots. \end{aligned}$$

Hence $$\lim_{\rho=1} \frac{\rho - \bar{r}}{\sigma_{\bar{r}}} = \lim_{\rho=1} \frac{\frac{1}{2}\gamma^{\frac{1}{2}}}{\sqrt{\left(\pi - \frac{16}{3\pi}\right)}}, \text{ when } \gamma = 0 \quad \ldots\ldots\ldots\ldots \text{(xxxiv)},$$

$$= 0.$$

Or we conclude that for $n = 4$, $r_{\rho\sigma_1} = 0$. We thus see that for all the cases in which our limit to $(\rho - \bar{r})/\sigma_{\bar{r}}$ becomes imaginary a direct investigation shows that $r_{\rho\sigma_1} = 0$. We now return for a moment to the method by which our limit for $n > 5$ was obtained, and we ask: Is the expansion by Taylor's series of the hypergeometrical legitimate? The doubt which exists in the process is this: The differentials of the hypergeometrical series for the values of α, β (i.e. both $\frac{1}{2}$ or both 1) and $\gamma = \frac{1}{2}(n+1)$ will ultimately become indefinite for some one or other value of n. Can we therefore use Taylor's theorem at all? I think we can. For if we are considering the limit when $n = s$ and accordingly a certain differential coefficient becomes infinite, say $\dfrac{d^u F}{dx^u}$ when $x = 1$, we have really to consider what happens to

$\epsilon^u \dfrac{d^u F}{dx^u}$ in such a case when ϵ approaches zero, and we can always make ϵ approach zero by taking it equal to $\epsilon'(n-s)$, so that the factor $n-s$ disappears from the denominator. For our purposes we need only the first two terms of Taylor's expansion. Thus, if θ be a fraction between 0 and 1, we shall have:

$$F_1\left(\tfrac{1}{2}, \tfrac{1}{2}, \tfrac{1}{2}(n+1), 1-\epsilon\right)$$
$$= F_1 \times \left\{1 - \frac{1}{2(n-3)}\epsilon + \frac{9}{8(n-3)(n-5)}\left(1 - \frac{25}{2(n-7)}\epsilon^\theta\right)\epsilon^2\right\} \quad \ldots(\text{xxxv}),$$

where the term in θ expresses the remainder after two terms. It may be said that this will become infinite when $n=7$. But as we are going to make ϵ approach zero we can take it $=(n-7)\epsilon'$, where ϵ' is also infinitely small, and we reach precisely the same limit as before. It seems therefore that for $n>5$ our method of approaching the limit will apply. At the same time a full discussion of the subject by a more competent mathematician is desirable.

III. *On the Standard Deviation of the Regression Coefficient, $r\sigma_1/\sigma_2$, and of the Standard Deviation of Arrays, $\sigma_1\sqrt{1-r^2}$, in samples drawn from a normal population.*

As before the frequency surface is

$$z = z_0' e^{-\frac{1}{2}\left(\frac{\sigma_1}{s_1}\right)^2} e^{-\frac{1}{2}\left(\frac{\sigma_2}{s_2}\right)^2} \left(\frac{\sigma_1\sigma_2}{s_1 s_2}\right)^{n-2} e^{\gamma r}(1-r^2)^{\frac{n-4}{2}} \quad \ldots\ldots\ldots(\text{xxxvi});$$

we need to multiply this in succession by

$$r\sigma_1/\sigma_2, \quad (r\sigma_1/\sigma_2)^2, \quad \sigma_1\sqrt{1-r^2}, \text{ and } \sigma_1^2(1-r^2),$$

and integrate over the whole volume, i.e. $r=1$ to -1, $\sigma_1=0$ to ∞, $\sigma_2=0$ to ∞.

We can do this most expeditiously by evaluating the integral of z multiplied by $r^p(1-r^2)^{\frac{1}{2}t}\sigma_1^{q_1}\sigma_2^{q_2}$.

$$W_{p,t,q_1,q_2} = \int_{-1}^{+1}\int_0^\infty\int_0^\infty z r^p \sigma_1^{q_1}\sigma_2^{q_2}(1-r^2)^{\frac{1}{2}t}\,dr\,d\sigma_1\,d\sigma_2$$
$$= z_0' s_1^{q_1+1} s_2^{q_2+1}\int_0^\infty\int_0^\infty e^{-\frac{1}{2}(\sigma_1'^2+\sigma_2'^2)}\sigma_1'^{n+q_1-2}\sigma_2'^{n+q_2-2} \times \int_{-1}^{+1} e^{\gamma r} r^p (1-r^2)^{\frac{1}{2}(n+t-4)}\,dr\,d\sigma_1\,d\sigma_2$$
$$\ldots\ldots\ldots\ldots(\text{xxxvii}),$$

where for brevity σ_1' and σ_2' are written for σ_1/s_1 and σ_2/s_2. Let

$$I_{p,t} = \int_{-1}^{+1} e^{\gamma r} r^p (1-r^2)^{\frac{1}{2}(n+t-4)}\,dr \quad \ldots\ldots\ldots\ldots\ldots(\text{xxxviii}).$$

We have to distinguish p odd from p even.

(i) p odd. Take $r=\sin\theta$ and expand the exponential; only the odd powers of γ remain, and

$$I_{p,t} = 2\int_0^{\frac{\pi}{2}}\left(\gamma\sin^{p+1}\theta + \frac{\gamma^3}{3!}\sin^{p+3}\theta + \ldots\right)\cos^{n+t-3}\theta\,\delta\theta.$$

Put $x=\sin^2\theta$, and we have:

$$I_{p,t}=\int_0^1\left(\gamma x^{\frac{1}{2}p}+\frac{\gamma^3}{3!}x^{\frac{1}{2}(p+2)}+\ldots\right)(1-x)^{\frac{1}{2}(n+t-4)}\,dx$$

$$=\frac{\Gamma(\frac{1}{2}(p+2))\,\Gamma(\frac{1}{2}(n+t-2))}{\Gamma(\frac{1}{2}(p+n+t))}\left(\frac{\gamma}{1!}+\frac{\gamma^3}{3!}\frac{p+2}{p+n+t}\right.$$

$$\left.+\frac{\gamma^5}{5!}\frac{(p+2)(p+4)}{(p+n+t)(p+n+t+2)}+\ldots\right)\ldots\ldots\ldots\text{(xxxix)}.$$

(ii) p even. We easily find:

$$I_{p,t}=2\int_0^{\frac{\pi}{2}}\left(\sin^p\theta+\frac{\gamma^2}{2!}\sin^{p+2}\theta+\ldots\right)\cos^{n+t-3}\theta\,d\theta$$

$$=\int_0^1\left(x^{\frac{1}{2}(p-1)}\theta+\frac{\gamma^2}{2!}x^{\frac{1}{2}(p+1)}\theta+\ldots\right)(1-x)^{\frac{n+t-4}{2}}\,dx$$

$$=\frac{\Gamma(\frac{1}{2}(p+1))\,\Gamma(\frac{1}{2}(n+t-2))}{\Gamma(\frac{1}{2}(p+n+t-1))}\left(1+\frac{\gamma^2}{2!}\frac{p+1}{p+n+t-1}\right.$$

$$\left.+\frac{\gamma^4}{4!}\frac{(p+1)(p+3)}{(p+n+t-1)(p+n+t+1)}+\ldots\right)\ldots\ldots\text{(xl)}.$$

Now $\gamma=\rho\sigma_1\sigma_2/(s_1s_2)=\rho\sigma_1'\sigma_2'$. Hence when we come to integrate these two expressions for σ_1' and σ_2' we require the integral, in the case of γ^u, of

$$U=\rho^u\int_0^\infty\int_0^\infty e^{-\frac{1}{2}\sigma_1'^2}e^{-\frac{1}{2}\sigma_2'^2}\,\sigma_1'^{n+u+q_1-2}\,\sigma_2'^{n+u+q_2-2}\,d\sigma_1'd\sigma_2'.$$

Changing the integration to y_1 and y_2 by the relations $\sigma_1'^2=2y_1$, $\sigma_2'^2=2y_2$, we have

$$U=\rho^u 2^{n+u+\frac{1}{2}(q_1+q_2)-3}\int_0^\infty e^{-y_1}y_1^{\frac{1}{2}(n+u+q_1-3)}\,dy_1\times\int_0^\infty e^{-y_2}y_2^{\frac{1}{2}(n+u+q_2-3)}\,dy_2,$$

or
$$U=\rho^u 2^{n+u+\frac{1}{2}(q_1+q_2)-3}\,\Gamma(\tfrac{1}{2}(n+u+q_1-1))\,\Gamma(\tfrac{1}{2}(n+u+q_2-1))\quad\ldots\text{(xli)}.$$

We have now to apply (xxxix) to (xli) to our special cases.

(*a*) *Mean value of* $r\sigma_1/\sigma_2=\{r\sigma_1/\sigma_2\}$, *say.*

Here $p=1$, $q_1=1$, $q_2=-1$, $t=0$. We have:

$$N\{r\sigma_1/\sigma_2\}=z_0's_1^2\frac{\Gamma(\frac{3}{2})\,\Gamma(\frac{1}{2}(n-2))}{\Gamma(\frac{1}{2}(n+1))}2^{n-3}\,\Gamma(\tfrac{1}{2}(n+1))\,\Gamma(\tfrac{1}{2}(n-1))$$

$$\times\left(2\rho+\frac{\rho^3}{3!}\frac{3}{n+1}2^3\frac{n+1}{2}\frac{n-1}{2}+\frac{\rho^5}{5!}2^5\frac{3.5}{(n+1)(n+3)}\frac{n+1}{2}\frac{n+3}{2}\frac{n-1}{2}\frac{n+1}{2}+\ldots\right)$$

$$=z_0's_1^2\,2^{n-3}\sqrt{\pi}\,\Gamma(\tfrac{1}{2}(n-2))\,\Gamma(\tfrac{1}{2}(n-1))\frac{\rho}{(1-\rho^2)^{\frac{1}{2}(n-1)}}\ldots\ldots\ldots\text{(xlii)}.$$

Hence, using (xix),

$$\{r\sigma_1/\sigma_2\}=\rho s_1/s_2=\rho\Sigma_1/\Sigma_2\ldots\ldots\ldots\ldots\ldots\ldots\ldots\text{(xliii)},$$

which is the expression for the mean slope of the regression lines of all samples. But as r is *not* symmetrically distributed about the value ρ and the mean value of r, i.e. $\bar{r}$, is not ρ (see (xxiii)), we could not *a priori* assume that the mean value

of $(r\sigma_1/\sigma_2)$ would be $\rho\Sigma_1/\Sigma_2$, the slope of the regression line in the sampled population.

(*b*) *Mean value of* $(r\sigma_1/\sigma_2)^2 = \{(r\sigma_1/\sigma_2)^2\}$, *say*.

Here we must put $p=2$, $q_1=2$, $q_2=-2$, $t=0$. We have:

$$U=\rho^u 2^{n+u-3}\,\Gamma\left(\tfrac{1}{2}(n+u+1)\right)\Gamma\left(\tfrac{1}{2}(n+u-3)\right),$$

and

$$N\{(r\sigma_1/\sigma_2)^2\} = z_0'\frac{s_1^3}{s_2}\frac{\Gamma\left(\frac{3}{2}\right)\Gamma\left(\frac{1}{2}(n-2)\right)}{\Gamma\left(\frac{1}{2}(n+1)\right)}2^{n-3}\,\Gamma\left(\tfrac{1}{2}(n+1)\right)\Gamma\left(\tfrac{1}{2}(n-3)\right)$$
$$\times\left[1+\frac{3\rho^2}{1!}\frac{n-3}{2}+\frac{5\rho^4}{2!}\frac{n-3}{2}\frac{n-1}{2}\frac{7\rho^6}{3!}+\frac{n-3}{2}\frac{n-1}{2}\frac{n+1}{2}+\ldots\right].$$

But the series on the right in square brackets is

$$\frac{d}{d\rho}\left(\rho\,(1-\rho^2)^{-\frac{1}{2}(n-3)}\right)=\frac{1+(n-4)\,\rho^2}{(1-\rho^2)^{\frac{1}{2}(n-1)}}.$$

Thus, using (xix) again:

$$\{(r\sigma_1/\sigma_2)^2\} = \tfrac{1}{2}\left(\frac{s_1}{s_2}\right)^2\frac{\Gamma\left(\frac{1}{2}(n-3)\right)}{\Gamma\left(\frac{1}{2}(n-1)\right)}(1+(n-4)\,\rho^2) \quad \ldots\ldots\ldots\ldots\text{(xliv)},$$

and, using R to denote $r\sigma_1/\sigma_2$:

$$\sigma_R^2 = \{(r\sigma_1/\sigma_2)^2\} - \{r\sigma_1/\sigma_2\}^2$$
$$=\frac{\Sigma_1^2}{\Sigma_2^2}\left(\frac{\Gamma\left(\frac{1}{2}(n-3)\right)}{2\Gamma\left(\frac{1}{2}(n-1)\right)}(1+(n-4)\,\rho^2)-\rho^2\right)$$
$$=\frac{\Sigma_1^2}{\Sigma_2^2}\left(\frac{1+(n-4)\,\rho^2}{n-3}-\rho^2\right)=\frac{\Sigma_1^2}{\Sigma_2^2}\frac{1-\rho^2}{n-3}.$$

Accordingly
$$\sigma_R=\frac{1}{\sqrt{n-3}}\frac{\Sigma_1}{\Sigma_2}\sqrt{1-\rho^2} \quad \ldots\ldots\ldots\ldots\text{(xlv)}.$$

Now the value usually taken for large samples* is:

$$\sigma_R=\frac{1}{\sqrt{n}}\frac{\Sigma_1}{\Sigma_2}\sqrt{1-\rho^2} \quad \ldots\ldots\ldots\ldots\text{(xlvi)}.$$

Thus the error in applying this to small samples lies solely in using $\frac{1}{\sqrt{n}}$ for $\frac{1}{\sqrt{n-3}}$. For the case of $\rho=\cdot6$ and $n=10$ (see our p. 190), the true value of $\sigma_R=\cdot5040\,\breve{R}$, while (xlvi) gave $\cdot4216\,\breve{R}$ and the method of our p. 190, $\cdot4205\,\breve{R}$. The error therefore in applying (xlvi) to compute σ_R does not lie in using the wrong values $\{\delta\sigma_1\delta\sigma_2\}$, $\{\delta r\,\delta\sigma_1\}$, $\{\delta r\,\delta\sigma_2\}$ for small samples, but in neglecting terms in higher products in (xxviii). The formula (xlv) is as simple as that for large samples, and I have not considered it needful to provide a table. The value under the radical is somewhat unexpected and suggests that special treatment is required for $n=3$. In this case we have really to find the integral of $\frac{\sigma_2}{s_2}\,e^{\rho r\frac{\sigma_1\sigma_2}{s_1 s_2}}e^{-\frac{1}{2}\left(\frac{\sigma_2}{s_2}\right)^2}$ and

* *Phil. Trans.* Vol. 191 A, p. 245 or *Biometrika*, Vol. IX. p. 9.

the limits being 0 and ∞, we approach at the lower limit the infinite area of a rectangular hyperbola. The fact is that for $n=3$, the frequency of cases of $\sigma_2=0$ is finite, and the second moment of the regression coefficient becomes infinite. The case is still more exaggerated for $n=2$. Hence we conclude that it is not possible to determine the mean slope of the regression line for samples of less size than four.

(c) *Mean value of* $\sigma_1\sqrt{1-r^2}=\{\sigma_a\}$, *say.*

Here $p=0$, $q_1=1$, $q_2=0$, $t=1$,

$$U=\rho^u\,2^{n+u-3+\frac{1}{2}}\,\Gamma(\tfrac{1}{2}(n+u))\,\Gamma(\tfrac{1}{2}(n+u-1)),$$

and we find:

$$N\sigma_a=z_0's_1^2s_2\,\frac{\Gamma(\frac{1}{2})\,\Gamma(\frac{1}{2}(n-1))}{\Gamma(\frac{1}{2}n)}\sqrt{2}\,2^{n-3}\,\Gamma(\tfrac{1}{2}n)\,\Gamma(\tfrac{1}{2}(n-1))$$
$$\times\left(1+\frac{\rho^2}{1!}\,\frac{n-1}{2}+\frac{\rho^4}{2!}\,\frac{n-1}{2}\,\frac{n+1}{2}+\frac{\rho^6}{3!}\,\frac{n-1}{2}\,\frac{n+1}{2}\,\frac{n+3}{2}+\ldots\right),$$

and this series is again $(1-\rho^2)^{-\frac{1}{2}(n-1)}$.

Hence using (xix) we find:

$$\{\sigma_a\}=\sqrt{2}\,s_1\,\Gamma(\tfrac{1}{2}(n-1)/\Gamma(\tfrac{1}{2}(n-2))$$
$$=\Sigma_1\sqrt{1-\rho^2}\,\frac{\Gamma(\frac{1}{2}(n-1))}{(n/2)^{\frac{1}{2}}\,\Gamma(\frac{1}{2}(n-2))}$$
$$=\Sigma_1\sqrt{1-\rho^2}\,\lambda_{n-1}/\sqrt{\tfrac{1}{2}n}\quad\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\text{(xlvii)},$$

in our previous notation. Thus the mean value of the standard deviation of arrays is *not* the value in the sampled population, until n is so large that we can take $\lambda_{n-1}/\sqrt{\frac{1}{2}n}$ as substantially unity. In this matter the mean of the array standard deviation differs from the mean of the slope of the regression line. I give a table by which the approach of $\{\sigma_a\}$ to σ_A, the standard deviation of arrays in the sampled population, can be judged.

TABLE III.

Ratio of Mean S.D. of Arrays in Samples to Array S.D. in Sampled Population.

Size of Sample	Ratio $\{\sigma_a\}/\sigma_A$	Size of Sample	Ratio $\{\sigma_a\}/\sigma_A$	Size of Sample	Ratio $\{\sigma_a\}/\sigma_A$	Size of Sample	Ratio $\{\sigma_a\}/\sigma_A$
2	·0000	*9*	·8511	*16*	·9189	*23*	·9442
3	·4607	*10*	·8670	*17*	·9238	*24*	·9466
4	·6267	*11*	·8798	*18*	·9282	*25*	·9488
5	·7136	*12*	·8904	*19*	·9321	*50*	·9747
6	·7675	*13*	·8992	*20*	·9356	*100*	·9874
7	·8042	*14*	·9068	*21*	·9388	*400*	·9969
8	·8308	*15*	·9132	*22*	·9416		

Clearly in this case we can hardly speak of "small samples" ceasing—for practical purposes even—before $n=50$ to 100.

(*d*) *Mean value of* $\sigma_1^2(1-r^2)$ *and standard deviation of the standard deviation of arrays in samples.*

Here $p=0$, $q_1=2$, $q_2=0$, $t=2$; and

$$U=\rho^u\,2^{n+u-2}\,\Gamma\left(\tfrac{1}{2}(n+u+1)\right)\Gamma\left(\tfrac{1}{2}(n+u-1)\right).$$

Hence
$$N\left\{\sigma_1^2(1-r^2)\right\}=z_0'\,s_1^3 s_2\sqrt{\pi}\,2^{n-2}\,\Gamma\left(\tfrac{1}{2}n\right)\Gamma\left(\tfrac{1}{2}(n-1)\right)$$
$$\times\left(1+\frac{\rho^2}{1!}\frac{n-1}{2}+\frac{\rho^4}{2!}\frac{n-1}{2}\frac{n+1}{2}+\frac{\rho^6}{3!}\frac{n-1}{2}\frac{n+1}{2}\frac{n+3}{2}+\ldots\right),$$

and the series $=(1-\rho^2)^{-\frac{1}{2}(n-1)}$.

We have using (xix):

$$\left\{\sigma_1^2(1-r^2)\right\}=s_1^2\,\frac{2\,\Gamma\left(\tfrac{1}{2}n\right)\Gamma\left(\tfrac{1}{2}(n-1)\right)}{\Gamma\left(\tfrac{1}{2}(n-2)\right)\Gamma\left(\tfrac{1}{2}(n-1)\right)}$$
$$=s_1^2(n-2),$$

or
$$\left\{\sigma_1^2(1-r^2)\right\}=\Sigma_1^2(1-\rho^2)\,\frac{n-2}{n}\qquad\ldots\ldots\ldots\ldots\ldots\ldots\text{(xlviii)}.$$

Thus the mean variance of the arrays in samples is not the variance in arrays of the sampled population, but the latter multiplied by the factor $(n-2)/n$.

To find the standard deviation of the standard deviation of arrays, we have:

$$\sigma^2_{\sigma_a}=\left\{\sigma_1^2(1-r^2)\right\}-\left\{\sigma_1\sqrt{1-r^2}\right\}^2$$
$$=\Sigma_1^2(1-\rho^2)\cdot\left(\frac{n-2}{n}-\frac{2\lambda^2_{n-1}}{n}\right),$$

or
$$\sigma_{\sigma_a}=\frac{\Sigma_1\sqrt{1-\rho^2}}{\sqrt{2n}}\,(2n-4-4\lambda^2_{n-1})^{\frac{1}{2}}\qquad\ldots\ldots\ldots\ldots\ldots\ldots\text{(xlix)}.$$

Clearly $\Sigma_1\sqrt{1-\rho^2}/\sqrt{2n}$ is the value we should have given to σ_{σ_a} on the theory of large samples. Calling this value $(\sigma_{\sigma_a})_L$ I have computed the values of the ratio $\sigma_{\sigma_a}/(\sigma_{\sigma_a})_L$ and they are given in Table IV.

TABLE IV.

Values of the Ratio $\sigma_{\sigma_a}/(\sigma_{\sigma_a})_L$ *for Various Sized Samples.*

Size of Sample	Ratio	Size of Sample	Ratio	Size of Sample	Ratio	Size of Sample	Ratio
2	·0000	*9*	·9808	*16*	·9907	*23*	·9939
3	·8525	*10*	·9834	*17*	·9914	*24*	·9942
4	·9265	*11*	·9853	*18*	·9919	*25*	·9944
5	·9524	*12*	·9868	*19*	·9924	*50*	·9974
6	·9651	*13*	·9881	*20*	·9928	*100*	·9987
7	·9725	*14*	·9891	*21*	·9932	*400*	·9997
8	·9774	*15*	·9900	*22*	·9936		

It is clear that for most practical purposes we may suppose that for a sample of 25 our ratio may be taken as unity.

It must be confessed of course that the results worked out in this paper only cover a limited series of sampled populations, i.e. those which follow the normal law. Their value is therefore rather for the light they throw on the general nature of small samples than on their direct application to any given series of small samples. We are a long way yet from a determination of the variation limits in the case of bi-variate small samples from *any* indefinitely large population whatever its law of distribution. Even in the present cases it is seen that the frequency distributions of the constants in samples have very complicated surfaces* of which it is not easy to get a clear geometrical conception.

For much help in the calculation of the tables I have to thank my colleagues Miss Ethel M. Elderton and Miss M. Moul; for computing the λ_n functions I am indebted to Miss B. N. Stoessiger, a statistical student in the Laboratory. Without their assistance the publication of this paper would have been indefinitely delayed, and I am most grateful for the aid they have given in the matter.

* There is some hope of expressing by a finite function the distribution of regression coefficients. But in this case the higher moments of the distribution will ultimately become infinite for small samples.

[While preparing this memoir from my notes, I received a paper from Professor Romanovsky. Writing without knowledge of the papers in *Biometrika* by "Student," Fisher, myself and others (Vol. VI. pp. 1—25, Vol. X. pp. 507—521, Vol. X. pp. 522—529, and Vol. XI. pp. 328—417) and naturally without knowledge of my present paper, Professor Romanovsky had reached, dealing only with the algebraical side, many of the published results and certain additional ones. While willing to publish the latter, the present cost of printing prohibited the reproduction of much work already published or about to be published in this *Journal*. Recognizing the difficulty of an Editor in such matters, I sent him a proof of this paper and asked him to cable if he were willing that I should add under the title his name to my own. I have at last received a letter from Professor Romanovsky. He is satisfied with the statement that many results contained in the present and earlier papers have also been obtained by him quite independently and by a different method. I trust for the sake of his additional results that his paper may shortly be published elsewhere. This note explains the delay in the issue of the present parts of *Biometrika*. K. P. *June* 25, 1925.]

Values of the Differences of the Powers of Zero.

Table of $q(p, s) = \dfrac{\Delta^p 0^{p+s}}{\Gamma(p+s+1)}$ *from* $p = 1$ *to* 10 *and* $s = 0$ *to* 10.

Values of p.

s	1	2	3	4	5	6	7	8	9	10	s
0	1·000,000,000,000	1·000,000,000,000	1·000,000,000,000	1·000,000,000,000	1·000,000,000,000	1·000,000,000,000	1·000,000,000,000	1·000,000,000,000	1·000,000,000,000	1·000,000,000,000	0
1	·500,000,000,000	1·000,000,000,000	1·500,000,000,000	2·000,000,000,000	2·500,000,000,000	3·000,000,000,000	3·500,000,000,000	4·000,000,000,000	4·500,000,000,000	5·000,000,000,000	1
2	·1666,66,666,667	·583,333,233,333	1·250,000,000,000	2·166,666,666,667	3·333,333,333,333	4·750,000,000,000	6·416,666,666,667	8·333,333,333,333	10·500,000,000,000	12·916,666,666,667	2
3	·041,666,666,667	·250,000,000,000	·750,000,000,000	1·666,666,666,667	3·125,000,000,000	5·250,000,000,000	8·166,666,666,667	12·000,000,000,000	16·875,000,000,000	22·916,666,666,667	3
4	·008,333,333,333	·086,111,111,111	·358,333,333,333	1·012,500,000,000	2·298,611,111,111	4·529,166,666,667	8·079,166,666,667	13·386,111,111,111	20·950,000,000,000	31·333,333,333,333	4
5	·001,388,888,889	·025,000,000,000	·143,750,000,000	·513,888,888,889	1·406,250,000,000	3·237,500,000,000	6·601,388,888,889	12·300,000,000,000	21 375,000,000,000	35·138,888,888,889	5
6	·000,198,412,698	·006,299,603,175	·050,016,534,592	·225,562,169,312	·741,732,804,233	1·989,616,402,116	4·625,925,925,926	9·671,957,671,958	18·628,174,603,175	33·604,414,682,540	6
7	·000,024,801,587	·001,405,423,280	·015,426,587,302	·087,632,275,132	·345,568,783,969	1·077,777,777,778	2·851,851,851,852	6·679,365,079,365	14·235,491,071,429	28·141,121,031,746	7
8	·000,002,755,732	·000,281,635,802	·004,284,060,847	·030,638,778,660	·144,695,216,049	·523,916,997,355	1·575,358,796,296	4·127,361,937,831	9·721,510,416,667	21·034,795,249,118	8
9	·000,000,275,573	·000,051,256,614	·001,083,829,365	·009,760,802,469	·055,162,863,757	·231,631,944,441	·790,558,449,074	2·314,315,476,190	6·017,912,946,431	14·238,267,471,342	9
10	·000,000,025,052	·000,008,546,944	·000,252,086,841	·002,860,825,517	·019,341,229,758	·094,114,940,326	·364,277,277,988	1·190,485,668,525	3·414,504,607,081	8·826,386,039,207	10

Values of s.

This table was computed by K. P. and E. M. E. from the formula $q(p, s) = \dfrac{p}{p+s}\{q(p, s-1) + q(p-1, s)\}$ and checked by the formula

$$\Delta^p 0^{p+10} = p^{p+10} - p(p-1)^{p+10} + \frac{p(p-1)}{2!}(p-2)^{p+10} - \ldots + (-1)^r \frac{p!}{(p-r)!\,r!}(p-r)^{p+10} + \ldots .$$

It was calculated for a special investigation, but it is thought that it may be of value to other computers, and it is accordingly published here. There may be a doubt as to the twelfth decimal in the latter part of the table, but the tenth row checked properly, and isolated values tested here and there in the table also checked. As far as the computers are aware, no table has hitherto been published showing the ratio of the differences of zero to the factorial.

JOURNAL OF ANATOMY

(ORIGINALLY THE JOURNAL OF ANATOMY AND PHYSIOLOGY)

CONDUCTED, ON BEHALF OF THE ANATOMICAL SOCIETY OF GREAT BRITAIN AND IRELAND, BY

Professor EDWARD FAWCETT, University of Bristol
Professor J. P. HILL, University of London
Professor ARTHUR ROBINSON, University of Edinburgh
Professor E. BARCLAY-SMITH, University of London
Professor Sir ARTHUR KEITH, Royal College of Surgeons

LINCOLN'S-INN-FIELDS, LONDON, W.C. 2

ANNUAL SUBSCRIPTION 40/- POST FREE

VOL. LIX

CONTENTS OF PART III.—APRIL, 1925. PRICE 10s. NET

On the Autostylism of Dipnoi and Amphibia (with Plates I—XXVI). By F. H. EDGEWORTH, M.D.

The Anomalous Right Subclavian Artery Considered in the Light of Recent Findings in Arterial Development; with a Note on Two Cases of an Unusual Relation of the Innominate Artery to the Trachea (with eleven text-figures). By JOHN CAMERON, M.B., CH.B.

Roentological Injection Masses—Old and New (with Plates I—IV). By FERGUS R. FERGUSON, M.D.

The Development of the Columella auris in the Crocodilia (with ten text-figures). By F. GOLDBY, B.A.

Note on a Right Lung which resembled a Left Lung in Presenting only Apical and Basal Lobes (with two text-figures). By DAVID HEPBURN, C.M.G., M.D.

Modern British Pelves in the "Alfred Hughes" Museum of Anatomy, University College, Cardiff. By DAVID HEPBURN, C.M.G., M.D.

An Unusual Variation of the Omo-Hyoid Muscle (with one text-figure). By J. TAYLOR, CH.M., M.D.

In Further Explanation of My Theory of the Polar and Trabecular Cartilages. By EDWARD PHELPS ALLIS, JR.

Obituary Notice. The late Emeritus Professor NICHOLAS KULCHITSKY (with portrait).

In Memoriam. JOHN IRVINE HUNTER, M.D., 1898–1924 (with portrait plate).

CAMBRIDGE UNIVERSITY PRESS
Fetter Lane, London, E.C. 4.

NOW READY

THE JOURNAL OF THE ROYAL ANTHROPOLOGICAL INSTITUTE

Vol. LIV. Part II. July to December, 1924

CONTENTS

La Race de Néanderthal et la race de Grimaldi ; leur rôle dans l'humanité. (The Huxley Memorial Lecture for 1924.) By Prof. R. VERNEAU. **Marital Gerontocracy in Africa.** By B. Z. SELIGMAN. **Neanderthal Man in Malta.** With an account of the survey of Dalam Cave (Ghar Dalam) by GEORGE SINCLAIR, A.M.I.C.E. (*With Plates XXXIII and XXXIV*). By Sir ARTHUR KEITH, F.R.S. **On the Palæolithic Deposits of Sawmills, Rhodesia.** By NEVILLE JONES. **Communications from the Anthropometric Laboratory of the University of Aberdeen.** By R. W. REID, M.D. (Aberd.), F.R.C.S. (Eng.), and JOHN H. MULLIGAN, M.B., Ch.B. (Aberd.). **South Indian Blow-Guns, Boomerangs and Crossbows.** (*With Plate XXXV.*) By JAMES HORNELL, F.L.S., F.R.A.I. **The Origin of Stencilling in the Fiji Islands.** (*With Plates XXXVI* and *XXXVII.*) By HENRY BALFOUR, M.A., F.R.S. **The Age of the Maya Calendar.** By RICHARD C. E. LONG, B.A. Miscellanea :—**Proceedings of the Royal Anthropological Institute. Further Note on the "Paccha" of Ancient Peru.** By T. A. JOYCE, M.A. Index. Contents, Vol. LIV.

PRICE 15/- NET

THE ROYAL ANTHROPOLOGICAL INSTITUTE
52, Upper Bedford Place, Russell Square, W.C.1.

General Agent:—FRANCIS EDWARDS, 83A, High Street, Marylebone, W. 1

OR THROUGH ANY BOOKSELLER

MAN

A MONTHLY RECORD OF ANTHROPOLOGICAL SCIENCE

Published under the direction of the Royal Anthropological Institute of Great Britain and Ireland.

Each number of **MAN** consists of at least 16 Imp. 8vo. pages, with illustrations in the text together with one full-page plate; and includes Original Articles, Notes, and Correspondence; Reviews and Summaries; Reports of Meetings; and Descriptive Notices of the Acquisitions of Museums and Private Collections.

Price, 2s. Monthly or £1 per Annum prepaid.

TO BE OBTAINED FROM THE

ROYAL ANTHROPOLOGICAL INSTITUTE,
52, Upper Bedford Place, W.C. 1.

AND THROUGH ALL BOOKSELLERS

(*All Rights reserved*)

BIOMETRIKA. Vol. XVII, Parts I and II

CONTENTS

PAGE

"Biometrika" Portrait Series, No. I. Abraham De Moivre . *Frontispiece*

I. A Study of Egyptian Craniology from Prehistoric to Roman Times. By G. M. Morant, M.Sc., Crewdson-Benington Student in Craniometry. (With eight figures in the text, folding table of measurements and six tissues in pocket) 1

Note.—The Definition of the Alveolar Point. By Karl Pearson, F.R.S. . 53

II. On the Moments of the Hypergeometrical Series. By Professor V. Romanovsky, Taskend 57

III. Probability Integrals for a Hypergeometric Series. By Professor Burton H. Camp, Ph.D. 61

IV. Determination of $\int_0^\theta \cos^{n+1}\theta\, d\theta$ for Large Values of n, and its Application to the Probability Integral of Symmetrical Frequency Curves. By John Wishart, M.A., B.Sc. (With one figure in the text) 68

V. On the Moments of the Distribution of Squared Standard-Deviations for Samples of N drawn from an Indefinitely Large Population. By A. E. R. Church, M.A., B.Sc. 79

VI. The Corpuscle Problem. A Mathematical Study of a Biometric Problem. By S. D. Wicksell, Lund. 84

VII. The Further Theory of Francis Galton's Individual Difference Problem. By J. O. Irwin, M.A., M.Sc. (With three figures and five diagrams in the text) 100

VIII. On the Multiple Correlation of Brothers, being a Note on Mr J. O. Irwin's Memoir, and on my Statement of the Application of Galton's Difference Problem to the Determination of the Degree of Relationship of Brothers, made in August 1902. By Karl Pearson, F.R.S. 129

IX. Heights and Weights of Patients in Mental Hospitals. By Rosa Greenwood, Cecily M. Thompson and Hilda M. Woods. (With one diagram in the text) 142

X. Further Contributions to the Theory of Small Samples. By Karl Pearson, F.R.S. 176

Table of the Differences of the Powers of Zero 200

Miscellanea:

(i). Some further Notes on Cancer and Goitre Distributions. By Percy Stocks, M.D. (With two maps in the text) 159

(ii). Recurrence Formulae for the Moments of the Point Binomial. By Ragnar Frisch, Kristiania 165

(iii). Review: Newsholme's The Elements of Vital Statistics in their bearing on Social and Public Health Problems. New Edition 171

(iv). Lebenserwartung und mittleres Alter der Lebenden. Von E. I. Gumbel, Heidelberg 173

The publication of a paper in *Biometrika* marks that in the Editors' opinion it contains either in method or material something of interest to Biometricians. But the Editors desire it to be distinctly understood that such publication does not mark assent to the arguments used or to the conclusions drawn in the paper.

A volume of *Biometrika* containing about 400 pages, with plates and tables, is issued annually.

Papers for publication and books and offprints for notice should be sent to Professor Karl Pearson, University College, London. It is a condition of publication in *Biometrika* that the paper shall not already have been issued elsewhere, and will not be reprinted without leave of the Editors. It is very desirable that a copy of all measurements made, not necessarily for publication, should accompany each manuscript. In all cases the papers themselves should contain not only the calculated constants, but the distributions from which they have been deduced. Diagrams and drawings should be sent in a state suitable for direct photographic reproduction, and if on decimal paper it should be *blue* ruled, and the lettering only pencilled.

Papers will be accepted in French, Italian or German. In the last case the manuscript should be in Roman not German characters. Russian contributors may use Russian but their papers will be translated into English before publication.

Contributors receive 25 copies of their papers free. Fifty additional copies may be had on payment of 17/- per sheet of eight pages, or part of a sheet of eight pages, with an extra charge for Plates; these should be ordered when the final proof is returned.

The subscription price, *payable in advance*, is 44*s. net* per volume (including packing and postage): single numbers 16*s. net* (including postage). Owing to the scarcity of early volumes, the following rates must now be charged for complete sets. Vols. I to XVI: Bound in buckram £56, in wrappers £50 net. Recent volumes may still be obtained at wrapper prices. Standard buckram cases with Darwin block, price 3*s.* 6*d.* + 6*d.* postage per volume. Index to Vols. I to V, 2*s. net*. Index to Vols. I to XV, 7*s.* 6*d. net*. Cheques must be made payable to Professor Karl Pearson and sent to the Secretary, Biometric Laboratory, University College, London, W.C. 1, to whom all orders for series and single copies should be addressed. All cheques should be crossed "*Biometrika Account*."

PRINTED IN ENGLAND BY W. LEWIS AT THE CAMBRIDGE UNIVERSITY PRESS.

Biometrika, Vol. XVIII, 1926

Subscribers are reminded that payment for this volume is now due and unless this reaches the publishers by March 31, 1926, the Journal cannot be obtained at 44*s.*, but only at the wrapper price of 64*s.* The publishers desire to remind Librarians and others that they cannot be responsible for the failure of agents to renew subscriptions before the above date. The Journal can always be obtained directly from the Biometric Laboratory, University College, London. Foreign subscribers must draw their cheques in sterling and if unstamped they must be drawn for £2. 4*s.* 2*d.* and not £2. 4*s.* 0*d.*